مبادئ
الأساليب الإحصائية

تأليف
الدكتور عبد العزيز فهمي هيكل

استاذ الاحصاء
بالمركز الدولي لتعليم الاحصاء
في بيروت

الطبعة الاولى
١٩٦٦

بيروت - لبنان

مبادىء الأساليب الاحصائية

مُقَدِّمَة

أقدم هذا الكتاب الى الطـالب المبتدىء في دراسة علم الإحصاء ، وقد راعيت فيه أن أناقش بشيء من التفصيل النواحي العملية المختلفة التي تسبق التحليل الرياضي للبيانات الاحصائية ، حيث اني لاحظت خلال السنوات الطويلة التي قمت فيها بتدريس هذا الموضوع ان الطالب بسبب تركيز كتب الاحصاء على الناحية الرياضية لا يدرك أهمية النواحي العملية التي يتطلبها البحث الاحصائي . ولا شك ان عدم إدراك الطالب لأهمية هذه النواحي يُبعد عن ذهنه الهدف الأساسي من تدريسه علم الاحصاء. ان الهدف الاساسي من تدريس طالب التجارة او الزراعة او الاجتماع علم الاحصاء هو تزويده باسلوب في البحث العلمي يستطيع ان يعتمد عليه كلما أراد بحث مشكلة من المشاكل بحثاً علمياً صحيحاً . لعل في هذه المناقشة الموجزة للنواحي العملية للأسلوب الاحصائي في البحث العلمي ما يشجع الطالب على محاولة التوسع فيها والاستزادة منها بالرجوع الى مراجعها في اللغات الاجنبية المختلفة . ذلك لأن الدول العربية المختلفة في هذه المرحلة من تاريخها هي في أشد الحاجة الى جيل من المتخصصين في البحوث العلمية حتى تهتدي بهم في تخطيطها لشؤونها الاقتصادية والاجتماعية والعمرانية .

واني مدين بالشكر الوافر للاستاذ مصطفى كريـــديه مدير دار النهضة العربية الذي تحمل عني عبء طبع هذا الكتاب وإخراجه على احسن وجه.

بيروت في اول آب ١٩٦٦

عبد العزيز فهمي هيكل

الفصل الأول

طبيعة علم الاحصاء

كان الانسان ولا زال يعمل جاهداً في البحث عن حقائق الكون المحيط به وأسرار الحياة التي يجهلها . وكان في باديء الامر يعتمد في بحثه هذا على تأملاته الخاصة فتوصل بذلك إلى فلسفات مختلفة متضاربة كان يتأرجح بالنسبة لها بين اليقين تارة والشك تارة اخرى. ولا يجب ان نقلل من أهمية الافكار التي تضمنتها هذه الفلسفات بالرغم من تضاربها ، إذ انها أخرجت الانسان من حالة العجز الفكري التي كانت تسيطر عليه ومهدت بذلك الطريق الى البحث العلمي .

ولا شك ان الانسان لم يقتنع بالنتائج التي توصل اليها عن طريق تأملاته حيث لم يستطع أن يسبغ عليها صفة الحقيقة المجردة ، وبذلك بدأ يفكر في منهاج آخر يعتمد عليه في بحثه .ان الملاحظة العابرة لا يمكن أن تعتبر حقيقة علمية بالمعنى الصحيح مهما بلغت أهميتها إلا إذا ثبت بالبرهان المادي الملموس ما لا يدعو الى الشك فيها ، وكيف السبيل الى إثباتها بغير التجربة في العمل؟

وبذلك بدأ الانسان يعتمد على التجربة في العمل كمنهاج لبحثه عن الحقيقة. ولا يعني ذلك أن فقدت ملاحظة الانسان أهميتها ، إذ ظلت هي نقطة الانطلاق في كثير من الابحاث العلمية . ان ملاحظة الباحث لظاهرة وما

يتولد عن ذلك من إحساس بمشكلة تتطلب الحل او التفسير هي في الحقيقه الخطوة الاولى في البحث العلمي .

ولقد استطاع الانسان بمنهاجه الجديد - التجربة في العمل - أن يتوصل الى كنز هائل من الحقائق العلمية في مختلف مجالات العلوم الطبيعية، الأمر الذي ساعد على تقدم هذه العلوم وعلى صلابة الأسس التي تعتمد عليها . وقد أفاد الانسان كثيراً من هذا التقدم حيث أصبحت حياته اليومية تقوم بالدرجة الأولى على استخدام مختلف الأشياء التي توصل إلى صنعها عن طريق البحث العلمي .

وبالرغم من تقدم العلوم الطبيعية تقدماً ملحوظاً بقيت العلوم الاجتماعية متخلفة وراءها حيث ان مجال استخدام التجارب في المعمل في هذه العلوم مجال ضيق محدود، إذ من الصعب في كثير من الحالات أن نخضع سلوك الانسان لتجارب نجريها عليه في المعمل حيث يقف دون ذلك عقبات كثيرة ترتبط بعادات المجتمع وتقاليده والأنظمة القانونية التي يسير عليها . على ان التقدم الحضاري في كثير من الدول وخاصه بعد الحرب العالمية الثانية قضى على كثير من هذه العقبات وبذلك بدأ العلماء يقومون بأبحاث على سلوك الانسان مستخدمين في ذلك منهاج التجربة في المعمل ، وبالرغم من ذلك لا نستطيع أن نتفائل كثيراً في امكانية اتساع نطاق استخدام هذا المنهاج في ميدان العلوم الاجتماعية ، حيث لا توجد قواعد محددة ثابتة تحكم سلوك الانسان في كل مجتمع وفي كل زمن .

لم يقف الانسان مكتوف اليدين امام هذه المشكلة، إذ استطاع أن يهتدي إلى منهاج آخر يساعده في الكشف عن الحقائق الخاصة بسلوك الانسان في النواحي الاجتماعية والاقتصادية والسيكلوجية والتربوية، انه المنهاج الاحصائي . وبذلك عندما ندرس الاحصاء انما ندرس في الواقع منهاجاً من مناهج البحث العلمي. ولا يعني ذلك أن الاحصاء ليس علماً قائماً بذاته، فهو في الواقع علم له

قوانينه وقواعده الرياضية الخـــاصة به ولكن مجال تطبيقه هو في خدمة العلوم الأخرى .

على أن استخدام المنهاج الاحصائي في البحث العلمي لا يقتصر فقط على العـــلوم الاجتماعية ، بل يمتد ايضاً الى بعض النواحي في العلوم الطبيعية ، وبشكل عام نستطيع أن نقول انه في كل حالة لا تتوفر بالنسبة لها المعلومات الكاملة عن الظواهر موضوع البحث يمـــدنا المنهاج الاحصائي (المعاينة) بالقواعد والاساليب التي يمكن استخدامها للتوصل الى قرارات حكيمة يمكن أن نطمئن اليها . فإذا أردنا مثلاً أن نقرر اذا كان نوع معين من السماد أفضل من غيره لا يستطيع الباحث أن يقوم بالتجارب على جميع قطع الأرض الموجودة في الدولة وبذلك يضطر الى أخذ عينة من الأراضي لاجراء تجاربه ، الأمر الذي يجعله في حاجة إلى فهم اسلوب المعاينة حتى يكون بحثه علمياً صحيحـــاً . كذلك إذا أردنا أن نفهم انماط الاستهلاك في مجتمع معين لا يستطيع الباحث أن يقوم بجمع المعلومات من جميع أسر المجتمع ، فيكون بذلك مضطرا الى أخذ عينة من الاسر، كيف يستطيع اختيار هذه الأسر؟، كيف يمكن أن تكون العينة غير متحيزة ؟ ، كيف يمكن أن يقدر منها المقاييس العامة لانماط الاستهلاك في المجتمع ؟ ، كيف يستطيع أن يحسب أخطاء هذه التقديرات ؟ ،كل هذه الاسئلة وغيرها لا يمكن الاجابة عليها الا بفهم أسلوب المعاينة .

ولا يجب أن يتبادر الى الذهن ان المنهاج الاحصائي يمكن استخدامه في بحث أية ظاهرة مهما كان نوعهـــا ، ذلك لأن هذا المنهاج يبدأ أولاً بجمع المعلومات عن الظـــاهرة أو الظواهر موضوع البحث ، فاذا لم تكن هذه المعلومـــات هي نفسها عبارة عن ارقام أو يمكن تحويلها الى أرقام يتعذر بـــذلك استخدام المنهاج الاحصائي . الا ان ذلك لا يحد كثيراً من مجالات استخدام المنهاج الاحصائي إذ يستطيع الباحث أن يصيغ الاسئلة التي يجمع

بها المعلومات في شكل يساعده على اتباع هذا المنهاج ؛ فاذا أردنا مثلاً أن نبحث رأي المجتمع في سياسة الحكومة نحو التعليم وسألنا كل مستجوب عن رأيه في هذه السياسة فاننا بهذا السؤال سوف نحصل على إجابات يستحيل تحويلها إلى معلومات رقمية تساعدنا في اتباع المنهاج الاحصائي ، أما إذا كانت الأسئلة في صيغة أخرى ، مثلاً : هل توافق على سياسة الحكومة نحو تحديد سن الالتحاق بالمداس الابتدائية ؟ وجعلنا الاجابة بنعم أم لا ، فاننا بمثل هذا السؤال نستطيع أن نحصل على اجابات يمكن تحويلها الى أرقام حيث نستطيع أن نحدد عدد الموافقين على هذه السياسة وعدد غير الموافقين .

والاحصاء كمنهاج للبحث العلمي يساعد الانسان في اتخاذ قرارات حكيمة عند مواجهة عـــدم التأكد ولذلك ازدادت أهميته في العصر الحاضر زيادة كبيرة ، حيث نواجه عدم التأكد في حياتنا اليومية وفي حيـــاتنا العامة بشكل لم يسبق له مثيل اننا نناقش كل ما توارثناه عن آبائنا وأجدادنا من أفكار وتعاليم ونطلب إجابة علمية على كل ما نوجهه من أسئلة . ان أسئلة مثل : هل هناك علاقة بين التدخين والسرطان ؟ وهل هناك علاقة بين تغذية الطفل وذكاؤه؟ وهل هنالك اختلاف جوهري بين خصوبة المرأة المتعلمة وغير المتعلمة ؟ وهل فعلاً يتناقص الميل الحدي للاستهلاك للفرد كلما زاد دخله ؟ وهل توجد فعلاً علاقـــة عكسية بين سعر الفايدة والاستثمار ؟ وهل تتغير أنماط الاستهلاك بتغير مستويات الدخول وما نوع هذا التغير وكيف يمكن قياسه ؟ وهل تتناقص التكلفة الحدية لأية سلعة في بادىء الأمر عنـــد زيادة الانتاج ثم تتزايد بعد ذلك ؟ ... الخ . اننا نقرأ عن اجابات لمثل هذه الأسئلة في الكتب المتخصصة ولكنها في الغالب تكون اجابات ممثلة لآراء أصحابها واعتقاداتهم . الأمر الذي يفقدها الصفة العلمية ويجعلها بذلك غير قادرة على اقناعنا .وحتى نستطيع الاجابة على هذه الاسئلة وغيرها اجابة علمية صحيحة يجب ان نتبع المنهاج الاحصائي في بحثنا .

وبذلك أصبح المنهاج الاحصائي هو المنهاج الذي بغيره لا يقتنع إنسان العصر الحديث عندما يوجه سؤالا يتطلب الاجابة العلمية الصحيحة عليه. ان رجل الأعمال لم يعد يثق في نتائج أية سياسة يتبعها نحو تسويق بضاعته إذا لم يكن المنهاج الاحصائي مؤيداً لها، كما ان التخطيط الاقتصادي والاجتماعي لا يمكن أن يكون تخطيطاً علمياً صحيحاً ما لم تؤيده البيانات الاحصائية وما لم يستخدم المنهاج الاحصائي في تتبع تنفيذه وتقييم نتائجه.

انواع البحوث الاحصائية :

تبين لنا مما تقدم أن مجـــالات استخدام المنهاج الاحصائي واسعة جداً حيث يستخدم في جميع العلوم الاجتماعية منها والطبيعية وحيث يساعدنا في اتخاذ قرارات حكيمة عند مواجهة عدم التأكد، الا اننا نستطيع أن نفرق بين ثلاث أنواع من البحوث الاحصائية هي :

أولاً . البحوث الاحصائية الوصفية وفيها تجمع المعلومات عن ظاهرة أو ظواهر معينة لا لخدمة هدف بذاته محدد سلفا، وانما بقصد توفير البيانات التي يمكن أن تخدم أغراضاً متعددة لباحثين مختلفين فيما بعد، وفي الغالب تقوم الأجهزة الاحصائية العامة في الدولة بهذه البحوث الوصفية امـــا على فترات دورية كماهو الحال في تعدادات السكان والتعدادات الصناعية والزراعية والتجارية، أو على فترات غير دورية كما هو الحال في بحث ميزانية الاسرة. فإذا أخذنا تعداد السكان مثلا نلاحظ أن الجهاز الاحصائي المكلف به يقوم بجمع معلومات تمثل النواحي الديموجرافية والاجتماعية والاقتصادية للسكان وبذلك تكون مادة خام لاجراء البحوث التحليلية عليها من قبل الباحثين المهتمين بأية ناحية من هـــذه النواحي المختلفة . وفي التعداد الصناعي تجمع المعلومات عن القوة العاملة المستخدمة في الصناعة وعن الاجور المدفوعة لهـا وعن القيمة الصافية المضافة وعن الاستثمار الكلي والصافي في النشاط الصناعي

وغير ذلك من المعلومات التي تساعد في بحث أي موضوع من هذه الموضوعات بحثاً تحليلياً .

ثانياً : البحوث الاحصائية التحليلية – وهي التي تجمع فيها المعلومات التي تخدم هدف معين في ذهن الباحث أو التي تساعد في تفسير مشكلة معينة لاحظها الباحث أو لاختبار صحة فرض معين ؛ والباحث في هذا النوع من البحوث لا يقتصر فقط على جمع المعلومات التي تخدم هدفه وانما يقوم كذلك بتحليلها التحليل الذي يراه مناسباً لاستنتاج المقاييس والمعاملات التي يرغب في التوصـــل اليها . وعند ما نتكلم عن جمع المعلومات في هذا النوع من البحوث لا نقصد أن الباحث لا بد أن يقوم هو نفسه بدراسة استقصائية في الميدان لجمع المعلومات، وقد يضطر فعلا الى ذلك، الا انه في كثير من الحالات قد يجد المعلومات التي يريدها متوفرة في النشرات الاحصائية التي تصدرها الهيئات الاحصائية العامة في الدولة أو لدى المؤسسات والهيئات والمصالح التي يمكن أن تقوم هي نفسها بجمعها بحكم مقتضيات عملها أو في التقارير التي صدرت في الماضي عن أبحـاث شبيهة . إلا أن الباحث في مثل هذه الحالة قد يجد المعلومات التي يريدها مصنفة ومبوبة تبويباً تحكمت فيه اعتبارات إدارية أو فنية أو محاسبية بحيث لا تخدم غرضه تماماً فيضطر الى اعادة تنظيمها وتنسيقها في الشكل الذي يريده ، ولا بد أن يستعين في هذا العمل بالمنهاج الاحصائي بالاضافة الى استعانته به عند إجراء التحليل .

ثالثا : البحوث الاحصائية التجريبية – ويستخدم هذا النوع من البحوث في ميـــادين مختلفة كالزراعة والطب والصحة العامة والنواحي التربوية والاجتماعية والاقتصادية .

وفي هذا النوع من البحوث يقوم الباحث بالتحكم في الظروف التي تجمع المعلومات على أساسها ، كما انه يستطيع ان يقوم بتكرار جمع المعلومات اذا

اقتضى الأمر إلى ذلك . وبذلك يختلف هذا النوع عن النوع السابق حيث انه في البحوث التحليلية لا يكون للباحث أية سيطرة على الظروف التي تجمع فيها المعلومات ، كما انه يضطر الى جمعها وقت وقوعها أو في وقت لاحق . ففي دراسة عن الضياع الذي ينتج عن دوران العمل في الصناعة مثلا ، اما أن يتفق الباحث مع المؤسسات التي تجمع المعلومات منها على تسجيل هذه المعلومات وقت وقوعها أو يقوم في وقت لاحق بجمعها . أما في البحوث التجريبية فان الباحث يواجه عدم التأكد بالنسبة لمشكلة ما ويريد أن يتوصل الى قرار يطمئن اليه ، وبذلك يضطر الى اجراء تجربة قد يعيد إجراءها مرارا قبل التوصل الى القرار الذي يريده . فاذا توصل باحث مثلاً ، إلى نوع جديد من التطعيم ضد مرض معين ويواجه مشكلة الاجابة على السؤال – هل هناك فرق جوهري بين التطعيم بهذا النوع الجديد والتطعيم بالنوع القديم ؟ – للإجابة على هذا السؤال نضطر الى اجراء تجربة على عينتين من البشر تطعم الأولى بالنوع الجديد وتطعم الثانية بالنوع القديم . وحتى يكون الفرق في النتائج راجع فقط الى نوع التطعيم المستخدم لا بد ان تتشابه العينتان في كافة الوجوه ولا بد أن يخضعها لنفس التأثيرات ، وهو مبدأ أساسي في إجراء التجارب وهو ان تكون وحدات التجربة متساوية تماماً في كل العوامل التي يمكن ان تؤثر عليها ولا تختلف الا في العامل الذي يراد معرفة تأثيره حتى يمكن استبعاد أثر العوامل الاخرى غير هذا العامل .

وبذلك نلاحظ ان الباحث في تجربته يكون في حاجة الى منهاج يسير عليه في تنظيم وتصميم التجربة حتى تكون النتائج علمية صحيحة . وقد اهتم علماء الاحصاء ببحث هذا التنظيم والتصميم ووضعوا القواعد والاصول بحيث أصبحت تكون سوياً فرعاً خاصاً في علم الاحصاء يسمى « تصميم التجارب ». لا شك انه في هذا النوع من البحوث يجري العمل بالتعاون من الباحث المختص بالمشكلة نفسها وبين الاحصائي وحتى يتحقق التعاون والتفاهم بينهما لا بد أن يكون كل منهما ملماً بعض الشيء باختصاص الآخر .

ما هو المنهاج الاحصائي :

تبينا ما تقدم أن الاحصاء علم قائم بذاته له قواعده وقوانينه الخاصة به ولكن أهميته تظهر في استخدامه كمنهاج للبحث في الميادين العلمية المختلفة . والسؤال الآن ما هي المراحل المختلفة التي يجب أن يمر بها العمل إذا أردنا استخدام هذا المنهاج ؟ ان المنهاج الاحصائي هو ، كما قدمنا ، منهاج علمي وليس منهاج فلسفي . ولذلك تكون أولى مراحله هي مرحلة جمع المعلومات التي تمثل واقع الظاهرة او الظواهر موضوع البحث حتى تكون المقاييس التي يمكن ان نتوصل اليها فيما بعد بالتحليل تابعة من هذا الواقـــع وليست مجرد تعبير عن رأي الباحث .

ولا شك ان اول مشكلة تصادف الباحث عند جمع المعلومات هي تحديد المصدر الذي يمكن أن يلجأ اليه . وهنالك نوعان من المصادر — المصادر التاريخية ومصادر الميدان — وتشمل المصادر التاريخية الوثائق والمطبوعات الحالية والقديمة . ومن أهم هذه المصادر النشرات الاحصائية الـــتي تصدرها الحكومات والهيئات الخاصة في الدول المختلفة . وتختلف هذه النشرات من دولة الى اخرى فهي قد تزيد بحيث يمكننا أن نجد فيها كل ما نريـــده من بيانات رقمية كما هو الحال في الدول المتقدمة احصائياً ، وقد تقل فلا نستطيع أن نجد الا اليسير من البيانات التي نريدها كما هو الحال في الدول المتخلفة من ناحية الوعي الاحصائي . ومن المصادر التاريخيـــة النشرات الاحصائيه التي تنشرها المنظمات المختلفة التابعة لهيئة الامم والنشرات التي يصدرها مكتب العمل الدولي فيما يختص بإحصاءات العمل .

ويمكن تقسيم المصادر التاريخيـــة الى نوعين : المصادر الاولية وهي التي تقوم بنشرها نفس الهيئة التي قامت بجمع المعلومات وتبويبها مثل جميـــع

النشرات التي تصدرها الادارات الاحصائية في الدول المختلفة ، والمصادر الثانوية وهي التي تقوم بنشرها هيئة غير التي قامت بجمع المعلومات وتبويبها، كأن يقوم شخص باقتباس بعض البيانات الموجودة في إحدى نشرات الادارة الاحصائية ويضعها كما هي أو بشيء من التحوير لتناسب الحالة التي يقوم بدراستها وذلك مثل الاحصاءات التي تنشرها الجرائد والمجلات والكتب .

وحينما يلجأ الباحث الى المصادر التاريخية للحصول على المعلومات التي يريدها يجب أن يأخذ في اعتباره ما يأتي :

١ – أن يلم بجميع التعاريف والمصطلحات التي تكمن وراء الأرقام التي تظهر في هذه النشرات ، حيث أن هذه التعاريف والمصطلحات يمكن أن تختلف من بحث الى آخر وبذلك يكون لها تأثيرها على معنوية الأرقام . فإذا كنا نقوم ببحث عن العمال المتعطلين مثلاً ؛ ولجأنا الى نشرة عن احصاءات العمال لا بد أن نتبين أولا تعريف المتعطل الذي استخدمته الهيئة التي قامت باعداد هذه الاحصاءات ، هل قامت باحصاء المتعطلين على أساس أن المتعطل هو كل شخص قادر على العمل وراغب فيه ومستعد لقبوله اذا عرض عليه ولا يجد عملا ، او على أساس أن المتعطل هو الذي سبق له العمل وتعطل . وغالباً تحتوي مقدمة النشرات الاحصائية على تفسير لجميع التعاريف والمصطلحات التي استخدمت في اعداد الاحصاءات التي تظهر فيها ، وإذا لم تكن واردة في المقدمة يمكن ان نبحث عنها أسفل الجداول في الجزء المخصص للملاحظات .

٢ – ان يراجع الاستمارة التي استخدمت في جمع المعلومات للتأكد من صحة تصميمها ودقة صياغة الاسئلة التي جاءت فيها

٣ – إذا كانت الاحصاءات التي تظهرها النشرات مصدرها دراسات

استقصائية بالمعاينة على الباحث ان يتثبت من أن العينة التي سحبت هي من النوع الذي يلائم المجتمع موضوع الدراسة ، ومن أن طريقة سحبها سليمة ، ومن جميع القوانين الرياضية التي استخدمت في تحديد حجم العينة والتقديرات منها وحساب اخطاء هذه التقديرات ، وبشكل عام يجب أن يتثبت من أن الدراسة اجريت وفقاً للقواعد الاحصائية السليمة المتعلقة بموضوع العينات . وعندما تكون الدراسة قد اجريت بالمعاينة فان التقرير النهائي عن الدراسة يحتوي غالباً على كل مـــا يريد الباحث أن يتثبت منه فيما يتعلق بعملية المعاينة نفسها .

٤ – إذا كانت الاحصاءات التي تظهرها النشرات مصدرها عمليات ميدان على الباحث أن يقف على المستوى التعليمي للعدادين الذين قـــاموا بجمع المعلومات وعلى نوع ومدى التدريب الذي حصلوا عليه قبل قيامهم بالعمل وعلى التعليمات التي اعطيت لهم وعلى تنظيم العمل وخاصة بالنسبة لعمليات الاشراف على العدادين ، ومراجعة البيانات ، وجميع العمليات الاخرى التي تجري في المكتب بعد جمع المعلومات من ترميز وجدولة وطبع .

٥ – التأكد من الفترة الزمنية التي تتعلق بها الاحصاءات .

٦ – قراءة عناوين الجداول قراءة فاحصة للتأكد من أن الأرقام الموجودة فيها تتعلق بالدولة بأكملها أو باجزاء منها فقط ، ومن انها خاصة بالفترة الزمنية التي يظهرها عنوان النشرة او بفترة زمنية اخرى ، ومن الاسس التي صنفت تبعاً لها الأرقام ومن وحدات الارقام، ومن انها أرقام كاملة أو مدورة وإذا كانت مدورة فكم رقم محذوف منهـــا (مثلاً ارقام بالمليون أو بالألف ... الخ) .

٧ – قراءة جميع الملاحظات والتفسيرات الموجودة في أسفل الجداول التي يريدها الباحث .

٨ – قراءة عناوين الأعمدة والأسطر في كعب الجدول وفهمها فهماً جيداً.

٩ – أي رقم وأي عنوان في الجدول يظهر بجانبه اشارة معينة يجب الرجوع الى هذه الاشارة في أسفل الجدول وقراءة الملاحظات المكتوبة أمامها مرة ثانية للالمام بما هو وارد فيها .

١٠ – يجب دائماً مراجعة المجاميع والمتوسطات للتأكد منها فلا يجب أن يأخذها الباحث كما هي قبل التحقق منها . ويحسن في بعض الاحيان حساب بعض المتوسطات والمعدلات فهي تساعد في تقييم الأرقام الوارد في الجدول من ناحية الدقة والشمول ، مثلاً يساعدنا حساب المعدلات الحيوية (معدلات المواليد والوفيات ...) في الكشف عن مدى شمول ارقام المواليد والوفيات في مناطق الدولة المختلفة ، كذلك يساعدنا حساب متوسطات أسعار السلع المختلفة في التحقق من ارقام القيم والكميات الواردة في جداول احصاءات التجارة الخارجية .

١١ – يجب قراءة الجدول بوعي وادراك شامل لمحتوياته حتى يستطيع أن نتنبه الى أي شيء يبدو غريباً في الجدول فنحاول ادراك هذه الغرابة بالرجوع الى التفسيرات والتوضيحات في مقدمة النشرة او في اسفل الجدول . أو في التصحيحات الموجودة في نهاية النشرة .

وبذلك يتبين لنا انه من الأفضل الالتجاء الى المصادر الاولية حيث انها تحتوي في الغالب على التفسيرات والتوضيحات المطلوبة وعلى جمع المستندات الخاصة بعملية جمع المعلومات مثل الاستمارة التي استخدمت والتعليمات التي اعطيت للعدادين وكذلك على وصف شامل لجميع خطوات العمل ولا يعني ذلك ان المصادر الثانوية لا تحتوي على مثل هذه الأشياء ، ففي بعض الأحيان تنقلها هذه المصادر من مصادرها الاولية ، الا انه غالباً ما يؤدي النقل الى اخطاء في الارقام او في التفسيرات. ومن جهة اخرى قد يكون من الأفضل الالتجاء الى المصادر الثانوية اذا كانت هذه المصادر اعادت نشر الاحصاءات في

شكل يكون اكثر فائدة للبــــاحث في بحثه وفي هذه الحالة تتوقف اهمية المعلومات التي تظهر فيها على مدى كفاية القائمين بهذا العمل وخبرتهم في موضوع البحث .

وعندما لا تتوفر المصادر التاريخية او عندما لا تكون كافية لاجراء البحث يكون الباحث مضطراً الى الالتجاء الى الميدان للحصول على المعلومات التي يريدها ، وفي هذة الحالة يسمى البحث ببحث ميدان – ويقوم هذا النوع من البحوث اساساً على تصميم استمارة تحتوي على اسئلة توجه الى المستجوبين المعنيين للحصول على اجاباتهم التي تكون هي بمثابة المعلومات التي يريدها الباحث . ويواجه الباحث في هذا النوع من البحوث مشاكل وصعوبات جمة ، الأمر الذي يجعلها في حــــاجة الى تنظيم محكم دقيق لجميع مراحل العمل الميداني والمكتبي وذلك لضمان دقة المعلومات التي يمكن التوصل اليها و ضمان انجاز العمل في الوقت المناسب ، واذا تذكرنا ان جمع المعلومات بهذه الطريقة يؤدي الى الاحتكاك بالجمهور او المنشآت والهيئات المختلفة واذا تذكرنا ان الوعي الاحصائي غالباً ما يكون ضعيفاً بل انه يمكن ان يكون سلبياً في في بعض الأحوال حيث يواجه الباحث رفض غالبية المستجوبين الاجابة أو تعمدهم اعطاء الاجابات الخاطئة المضللة ، اذا تذكرنا ذلك ادركنا مدى مـــا يحتاجه الباحث من تنظيم للتغلب على هذه الصعوبة بالدرجة الأولى وعلى الصعوبات الاخرى التي يمكن أن تنشأ عن أوجه .النقص في جهاز الموظفين القائم بالعمـــــل . وفي الفصل التالي نناقش بشيء من التفصيل خطوات العمل المختلفة التي يتم وفقاً لها جمع المعلومات من الميدان .

بعد أن تتم مرحلة جمع المعلومات تبدأ المرحلة الثانية في المنهاج الاحصائي وهي الخاصة بعرض هذه المعلومات عرضاً يساعد فيما بعد على تحليلها .ويكون عرض المعلومات بواسطة تصنيفها وتبوينها في جداول ثم بتوضيحها بالرسومات البيانية كلما كان ذلك ممكناً . ولا شك أن عملية التصنيف والتبويب لا

تأخذ مجهوداً كبيراً من قبل الباحث الذي يكون مصدر بحثه المصادر التاريخية المختلفة . ولا يجب أن يتبادر الى ذهننا ان الباحث في مثل هـذا النوع من البحوث يأخذ المعلومات كما هي مبوبة في النشرات ، حيث انه في الغالب يضطر الى إعادة تنسيقها وتبويبها التنسيق والتبويب الذي يتفق والبحث الذي يقوم به . اما في بحوث الميدان فان المعلومات تأتي الى المكتب في استمارات يصل عددها في بعض التعدادات (تعداد السكان مثلاً) الى الملايين والمعلومات كما تظهر في الاستمارات لا يمكن بأي حال من الأحوال أن تساعد في فهم اتجـاه الظاهرة أو الظواهر موضوع للبحث ، بل لا يمكن أن يستوعبها الباحث في ذهنه مهما أُوتي من الذكاء وبذلك تكون عملية التصنيف والتبويب عملية واجبة حتى يمكن أن يأخذ المنهاج الاحصائي مجراه وحتى يمكن أن ينتهي بنا الى النتائج التي نقوم بالبحث من أجلها ، ذلك لاننا وان كنا نجمع المعلومات عن كل وحدة على حدة،الا انهذه المعلومات الفردية لا تهمنا في شيء حيث ان اهتمامنا يتجه الى معرفة الاتجاهات العامة للظواهر موضوع البحث . فاذا كنا نقوم بدراسة عن مستوى الاجور ، مثلاً ، فاننا قد نسأل كل عامل على حدة عـن أجره ، وسؤالنا هذا لا يعني اننا نرغب في معرفة أجر كل عامل فعلاً وانمـا لاننا نستطيع أن نتوصل عن طريق معرفة اجور العمال الى تحديد المتوسط العام لأجرهم وتوزيعهم في فئات معينة للأجر وغير ذلك من الحقائق التي نرغب في معرفتها عن ظاهرة الأجر بشكل عام . ولهذا عند تبويب الأجور في فئات معينة فان الاجور تفقد صفتها الشخصية وتصبح في نظرنا مجرد ظاهرة نجري عليها تحليلنا الذي نريده .

وبعد أن تتم مرحلة عرض المعلومات تنتهي بذلك البحوث الاحصائية الوصفية حيث لا يتبقى الا طبع النشرة التي تحتوي على الجداول والرسوم التي أمكن التوصل اليها . اما في البحوث التحليلية والتجريبية حيث يكون الهدف الأساسي هو التوصل الى المقاييس المختلفة التي تدل على اتجاهات الظواهر

موضوع البحث تكون المرحلة التي تتلو عرض المعلومات هي اجراء عمليات التحليل المختلفة التي يرغب الباحث في اجرائها . والتحليل الاحصائي يشمل عمليات كثيرة ندرس بعضاً منها في هذا الكتاب مثل حساب المتوسطات والتشتت والالتواء والارتباط والانحدار ... الخ

ويضيف بعض الكتاب إلى المراحل الثلاث السابقة مرحلة رابعة هي التفسير أي توضيح مدلولات المقاييس المختلفة التي أمكن التوصل اليها بالتحليل، والواقع ان التفسير وان كان ينبني على المقاييس التي يمكل التوصل اليها باستخدام المنهاج الاحصائي إلا انها عملية ليست ذات طبيعة احصائية بحتة حيث انها تحتاج الى خبرات لا تكون خبرات احصائية مجردة ، إذ ان تحديد مدلولات المقاييس يتوقف على المعرفة المتخصصة بالناحية العلمية التي يعمل فيها الباحث . فلو توصلنا مثلاً ، باستخدام المنهاج الاحصائي الى مقياس معين للارتباط بين لون الشعر ولون العينين فان ادراك هذا الارتباط وتفسيره ليس من واجب الاحصائي .

التضليل الاحصائي :

تبين لنا ان استخدام المنهاج الاحصائي يوصلنا الى نتائج رقمية ذات قوة اقناع لا حد لها حيث ان الارقام أصبحت تمثل في ذهن انسان العصر الحديث الحقيقة التي لا يمكن مناقشتها ، إلا اننا نلاحظ أن تفسير النتائج الرقمية التي نصل اليها عن طريق استخدام المنهاج الاحصائي يمكن أن يكون مضللاً ما لم نعالج هذا التفسير بحكمة ومهارة وما لم نكن موضوعيين في تعليقاتنا على هذه النتائج ، وليس ادل على ذلك من النقاش الذي يحتدم بين فريقين من الناس فيما يتعلق بموضوع معين وكل منهما يؤيد رأيه ببيانات احصائية قد تكون في بعض الاحيان مستقاة من نفس النشرة . والسؤال الآن ما هي الاسباب التي يمكن أن تؤدي الى ذلك ؟ هناك أسباب كثيرة نورد بعضها في الآتي :

١ – الجهـــل بالنعاريف والمصطلحات التي تكمن وراء الأرقام . فعند مقارنة البطالة في دولة ما في فترتين زمنيتين ، مثلاً ، يمكن أن نقع في خطأ فادح فنقول ان البطالة في هذه الدولة قد ازدادت حيث ان الأرقام تدل على ذلك مع اننا لو حاولنا ان ندرك ما وراء هذه الارقام وأدركناه ادراكاً صحيحاً لتوصلنا الى عكس النتيجة السابقة بالرغم من الدلالة الظـــاهرية للأرقام ، اذ قد تكون الارقام في الفترة الأولى مبنية على تعريف للمتعطل يقتصر على اعتبار الشخص متعطلاً اذا كان قـــد سبق له العمل وتعطل ، كما انها قد تكون مأخوذة من مكاتب التشغيل ونحن نعلم علم اليقين أن القانون لا يجبر كل متعطل على تسجيل اسمه في هذه المكاتب وبذلك يوجد عدد كبير من المتعطلين الذين لا يسجلوا اسمائهم اعتماداً على قدرتهم على أن يجدوا عملا لأنفسهم دون الألتجاء الى مكاتب التشغيل ، بينما في الفترة الثانية يمكن أن تكون الارقام مبنية على تعريف اكثر شمولاً حيث اعتبر الشخص متعطلاً سواء سبق له العمل او لم يسبق له ذلك طالما انه قادر على العمل ويبحث عنه دون أن يجده ، كما انها قد تكون نتيجة لدراسة استقصائية شاملة لكل فئات المجتمع ولذلك كله تبدو الارقام في هذه الفترة اكبر منها في الفترة الاولى .

كذلك يمكننا ان نناقش فنقول ان مستوى الاجور في دولة ما قد ارتفع خلال فترة معينة بينما يناقش فريق آخر من الناس فيقول ان العكس هو الصحيح ، وكل منا يؤيد رأيه ببيانات احصائيه نستقيها جميعاً من النشرة الاحصائية الرسمية . واذا اردنا ان ندرس الموضوع دراسة موضوعية امكننا ان نتبين ان رأي كل من الفريقين صحيح اذا اخذنا في الاعتبار نوع البيانات الاحصائية التي تؤيد الرأيين ـ فقد يكون الفريق الأول معتمدا في رأيه على بيانات خـــاصة بكسب العمل ويمكن ان تكون قد ارتفعت بسبب زيادة ساعات العمل الاضافي او بسبب زيادة الاجور العينية التي تمنحها المؤسسات للمستخدمين ، بينما يكون الفريق الثاني معتمدا في رأيه على بيانات خاصة

بمعدلات الاجور التى تكون قد انخفضت بسبب استخدام الآلات الحديثة التي تجعل العمل أكثر آلية واقل عناء .

٢ – عدم معرفة مزايا وعيوب الطرق المختلفة التي يمكن ان نستخدم في جمع المعلومات التي سوف يجري تحويلها الى احصاءات بالتصنيف والتبويب. فاذا كنا ندرس الحالة الصحية في دولة ما خلال فترة طويلة من الزمن فقد تفاجئنا الارقام بازدياد عدد الامراض التي تظهرهـا الجداول ، الأمر الذي يجعلنا نصل الى استنتاج خاطىء فنقول ان الحالة الصحية في هذه الدولة في تدهور مستمر . ولكننا اذا ادركنا ان مصدر الأرقام الخاصة بالامراض هو المستشفيات والمؤسسات الصحية المختلفة استطعنا أن نفهم سبب زيادة هذه الارقام حيث ان توسع الدولة في انشاء المستشفيات والمؤسسات الصحية في المناطق المختلفة يؤدي الى التبليغ عن امراض لم يكن من الممكن التبليغ عنها عندما كانت هذه المناطق محرومة من جميع الخدمات الصحية

كذلك يمكن أن يكون مصدر المعلومات الدراسة الاستقصائية بالمعاينة فاذا لم ندرك ان التحيز يمكن أن يتطرق الى مثل هذه الدراسات اذا لم تتبع الاصول الاحصائية السليمة ، فاننا لا بد وأن نقع في أخطـاء فادحة اذا اعتمدنا على نتائج هذه الدراسات دون أن نتأكد من اتباع هذه الاصول . وفي بعض الاحيان قد يكون التحيز متعمداً بقصد التضليل ، والتحيز المتعمد يمكن أن يكون واعياً ويمكن أن يكون لا شعوريا . ففي كثير من الدراسات التي يكون هدفها الدعاية يكون التحيز متعمداً والمثل على ذلك ما يقوم به بعض الجرائد والمؤسسات التليفزيونية من دراسات استقصائية لا ثبات اراء يؤيدونها سلفاً. الا انه في بعض الأحيان يكون الباحث متأثراً بإحساس معين تأثر باطنياً ، الأمر الذي يجعله في تصميم استمارة البحث يصيغ الاسئلة بشكل يعطيه اجابات معينة يؤيدها هو نفسه ، والمثل على ذلك دراسة استقصائية عن أسباب فشل الزواج يقوم بها رجل فشل في زواجه ويلقى لذلك كل التبعة على المرأة ، فانه لا شعورياً يعمل على إثبات هذه النتيجة بأية وسيلة .

٤ – عدم جمع المعلومات الكافية عن الظـــاهرة أو الظواهر موضوع البحث ، ففي دراسة عن الارتباط بين درجات الطلبة في مادة الاحصاء ودرجاتهم في مادة المحاسبة يمكن ان نصل الى نتيجة مضللة اذا اعتمدنا في دراستنا على عينة صغيرة من عشرة طلبة مثلاً حيث أنه في هذه الحالة يكون احتمال تأثير الصدفة على النتيجة احتمال كبير جداً . وفي بحث عن أفضلية دواء جديد في معالجة مرض معين لا يمكن أن نثق بالنتيجة اذا كانت مبنية على تجربة العقار على عدد قليل جداً من الأشخاص . وخلاصة القول إن نتائج الدراسات بالمعاينة تعتمد الى حد كبير على حجم العينة أي على عدد الوحدات التي تتكون منها العينة وسوف نتبين عند دراسة موضوع العينات ان هناك علاقة عكسية بين حجم العينة ومقدار الخطأ المحتمل حدوثه في النتائج . ولا يجب أن يتبادر الى الذهن ان مجرد تكبير حجم العينة فيه الضمان الكافي لدقة نتائج العينة ، فهناك قواعد وأصول كثيرة خاصة بهذا الموضوع التي لا بد من أتباعها لضمان دقـة تمثيل العينة للمجتمع الذي سحبت منه .

٥ – عدم معرفة مضمون الفئات المختلفة التي صنفت تبعا لها البيانات الاحصائية ويحدث هذا غالباً عند مقارنة أرقام التجارة الخارجية لدولتين أو لدولة واحدة في فترتين مختلفتين ، فمفهوم كلمة مواد أولية أي مكوناتها يمكن أن تختلف من دولة الى أخرى أو من وقت الى آخر في نفس الدولة ، ولعل ذلك هو السبب في قيام المكتب الإحصائي التابـع لهيئة الأمم بوضع تصنيف موحد للبضائع التي يجرى تبادلها فيما بين الدول المختلفة حتى تكون المقارنة بين احصاءاتها مقارنة صحيحة غير مضللة. كذلك يحدث هذا الخطأ عند مقارنة التوزيع النسبي للسكان تبعا للمهن أو تبعاً للنشاط الإقتصادي لاختلاف مفهوم الفئات المختلفة التي تدخل في تصنيف المهن أو في تصنيف النشاظ الإقتصادي، ولعل ذلك هو السبب كذلك في ان الهيئات الدولية المتخصصة (مكتب العمل الدولي ومكتب الإحصاء التابع لهيئة الأمم) قد وضعت تصنيفات موحدة وأوصت باستخدامها . كذلـك يحدث هذا الخطأ عند مقارنة احصاءات

الأمراض المختلفة حيث يمكن ان يختلف مفهوم الأمراض الصدرية أو الأمراض الباطنية من دولة الى اخرى أو من وقت الى آخر ولذلك وضعت منظمة الصحة العالمية تصنيفاً موحدا للامراض والايذاءات وأسباب الوفاة وأوصت باستخدامه في الدول الأعضاء حتى يمكن مقـــارنة احصاءاتها مقارنة علمية صحيحة . وقد بدأ كثير من الدول فعلاً في استخدام هذه التصنيفات الدولية.

٦ - كذلك يمكن أن نقع في أخطاء فادحة بسبب الخطأ في تفسير النتائج التي نصل اليها . ويحدث ذلك أما بسبب اهمال أحد العوامل التي يمكن أن يكون لها تأثير على الظاهرة والظواهر موضوع البحث أو بسبب عدم الالمام بموضوع البحث الماما كافياً . فاذا كنا نقوم بتجربة لاختبار أثر عامل معين لا بد وأن نجعل باقي العوامل الاخرى التي يمكن أن يكون لا تأثير على هذه الظاهرة مشتركة كما سبق أن قدمنا ، فاذا أهملنا أحد العوامـــل الاخرى وأبقيناه مختلفاً خلال اجراء التجربة فان النتيجة لا يمكن أن ترتبط بالعامل الذي ندرسه فقط . كذلك اذا كنا ندرس الارتبـــاط بين الاتجاه العام لسلسلتين زمنيتين ، مثلاً الارتباط بين الاتجاه العام لصادرات الدولة في عدة سنوات والرقم القياسي لاسعار الجملة في هذه السنوات ، ولم نخلص السلسلتين من أثر الانواع الاخرى من التغيرات التي يمكن أن تؤثر عليهما (التغيرات الموسمية والدورية والعرضية) فان النتيجة التي نصل اليهـــا لا يمكن أن تكون ذات دلالة صحيحة على وجود ارتباط بين الاتجاه العام للظاهرتين . كذلك اذا كنا نقارن مستوى الاجور في صناعتين مختلفتين وتبين من النتائج الرقمية أن المستوى في أحدهما أعلا من الآخر ، فاننا يمكن أن نقع في خطأ فادح إذا فسرنا ذلك نتيجة اختلاف الصناعتين ، حيث يمكن أن يكون السبب في الاختلاف ليس راجعاً الى اختلاف نوع الصناعة بقدر ما هو راجع إلى أن احدي الصناعتين تتركز في المدن حيث مستوى الاجور أعلى والاخرى تتركز في الريف حيث مستوى الاجور أقل . كذلك إذا كنا في صدد مقارنة مستوى المعيشة في دولة مـــا في فترتين زمنيتين واعتمدنا في ذلك على أرقام

الدخل الشهري فاننا يمكن ان نقع في اخطاء فادحة اذا لم نأخذ في اعتبارنا نمو السكان خلال الفترة موضوع البحث ، وتطور الاسعار خلال هذه الفترة ، واختلاف توزيع الدخل بين السكان حيث يمكن ان يكون التوزيع في الفترة الاولى اقل عدالة منه في الفترة الثانية .

تبين لنا من هذه المناقشة ان هناك أسباب كثيرة يمكن أن تجعل أحكامنا على الظواهر التي نبحثها احكاما خاطئة ، ويحدث ذلك غالباً عندما لا يكون الانسان ملماً بالمنهاج الاحصائي الماماً كافياً . ولذلك فدراستنا لهذا المنهاج سوف تساعدنا بالاضافة الى معرفة كيفية استخدامه في البحث العلمي ، في ألا نقع في الاخطاء الفادحة التي يمكن ان نقع فيها بسبب جهلنا به خاصة إذا كنا نقوم بابحاث في الميادين العلمية التي تهمنا .

الفصل الثاني

جمع المعلومات

في هذا الفصل اناقش الخطوات المختلفة التي يجب علينا أن نتبعهـــا إذا أردنا النزول الى الميدان لجمع المعلومات التي تلزمنا في البحث الذي نريد القيام بـــه . وقبل أن نبدأ هذه المناقشة يجب أن اشـير الى ان هناك كثير من الادارات الحكومية التي يتجمع لديها معلومات كثيرة في السجلات التي تمسكها فيما يختص بالاعمال التي تقوم بها ، وهي تقوم بذلك ليس لغرض احصائي وانما لأغراض اخرى يمكن أن تكون اغراض ادارية أو أغراض قانونية أو أغراض مشتركة بينهما ؟ فمديرية الجمارك مثلاً تسجل في دفاترها معلومـــات مختلفة عن كل بضاعة تخرج من الدولة أو تدخل اليهـــا ، ومديرية الاحوال الشخصية تقوم كذلك بتسجيل معلومات مختلفة عن كل مولود وكل وفاة وكل زواج وكل طلاق ، وأقسام البوليس تقوم ايضاً بتسجيل معلومات مختلفة عن كل حادث طريق يحدث في الدولة . الخ .. هذه المعلومات كما هي مسجلة في سجلات هذه المديريات المختلفة لا تعتبر احصاءات ، ذلك لانها معلومات فردية ، أي عن كل حالة على حدة ولكنها تعتبر مصدرا هاماً للاحصاءات الخاصة بهذه الموضوعات المختلفة وذلك يكون بتصنيفها وتبويبها في جداول تظهر ارقاماً مجمعة . والمهم في هذا الصدد ان المعلومات التي تسجلها هذه المديريات المختلفة قد تكون مسجلة بشكل لا يساعد على تحويلها الى احصاءات بالتصنيف والتبويب ، أو قد تكون غير كافية لاعداد احصاءات عن هذه

الموضوعات تساعد الباحث في بحث معين يقوم به . فاذا كنا نقوم ببحث عن خصوبة المرأة في فئات العمر المختلفة مثلاً ، نكون في حاجة الى بيان عن عمر الام لكل مولود فاذا لم تكن مديرية الأحوال الشخصية تهتم بتسجيل هذا البيان فاننا لا نستطيع أن نستمر في بحثنا حيث لن نجد المعلومات التي نريدها . وبذلك يكون من الأفضل لو تعاونت هذه المديريات المختلفة مع مديرية الاحصاء او مع الأقسام الاحصائية في الوزارات المختلفة على تصميم الاستمارات التي تجمع بواسطتها المعلومات الخاصة بهذه الموضوعات وبذلك يمكن تضمينها للأسئلة المختلفة التي تهم الباحثين في ابحاثهم ، كما يمكن تصميم هذه الاسئلة بالشكل الذي يساعد في تطبيق المنهاج الاحصائي .

كذلك نلاحظ ان خطوات العمل عندما يكون الميدان هو مصدر المعلومات التي نريد جمعها لا تختلف سواء أردنا ان نجمع المعلومات بالعد الشامل أي من جميع وحدات المجتمع الاحصائي [1] موضوع البحث او بالمعاينة أي من بعض وحدات هذا المجتمع الذي سحبت منه بحيث تكون مع بعضها صورة صادقة له ، والاختلاف الوحيد بالاضافة الى الخطوات الخاصة بالمعاينة والتي تقوم بها فقط عندما تكون الدراسة بهذه الطريقة ، هو في نطاق العمل حيث يكون النطاق واسع جداً في التعدادات (العد الشامل) وأضيق بكثير

---

(١) يقصد بالمجتمع الاحصائي مجموعة الوحدات موضوع البحث ويجب ان تكون هذه الوحدات معروفة بصورة واضحة بحيث يمكن تمييزها عن غيرها من الوحدات التي تكون مجتمعاً آخر . فاذا كنا بصدد بحث عن الموظفين المصنفين في لبنان ، يكون كل موظف مصنف في الحكومة وحدة احصائية ومجموع الموظفين المصنفين يكونوا المجتمع الاحصائي . واذا كنا بصدد دراسة عن الصناعة في لبنان فان مجموع المؤسسات الصناعية تكون هي المجتمع الاحصائي وكل مؤسسة منها هي الوحدة الاحصائية . أما اذا كنا بصدد دراسة عن صناعة الغزل والنسج فقط فان مجموع المؤسسات التي تعمل في هذا النشاط تكون هي المجتمع الاحصائي وكل مؤسسة منها هي الوحدة الاحصائية .

والمجتمعات الاحصائية تكون دائمة التغير من وقت الى آخر وهذا ما يجعل الدراسة الاحصائية لأي مجتمع مرتبطة بالوقت الذي جمعت فيه المعلومات الخاصة بهذه الدراسة .

في العينات . وقد يبدو هذا الاختلاف بسيطاً ، إلا أن ما يترتب على ذلك من ناحية الدقة في النتاج امر جوهري . ذلك لأنه عندما يكون العمل على أساس العـــد الشامل تزداد نسبة الخطأ بسبب صعوبة التنظيم والاشراف الدقيق ، بينما اذا كان العد (جمع المعلومات) بالمعاينة فان الجهاز القائم بالعمل يستطيع ان يتحكم في التنظيم والاشراف بحيث يمكن أن ينقص الاخطاء الى درجة كبيرة . ولعل ذلك هو السبب في ان كثيراً من الاحصائيين يعتقدون ان الدراسة بالمعاينة وخاصة بالنسبة للموضوعات المعقدة تأتي بنتائج أكثر دقة من الدراسة بالعد الشامل .

وتقودنا هذه المناقشة الى بحث موضوع الاخطاء التي يمكن أن تظهر في نتائج الابحاث الاحصائية والتي يجب أن يعمل الجهاز القائم بالعمل على تلافيها أو انقاصها الى ادنى حد ممكن ، ولن يتأتى له ذلك إلا بالتنظيم المحـــكم والاشراف الدقيق والمراجعة المستمرة لنتائج كل خطوة من خطوات العمل ، ومما لا شك فيه ان الدراسة بالمعاينة تهيأ لنا فرصة أفضل لتحقيق كل ذلك بسهولة ويسر .

ان أول نوع من الاخطاء التي يمكن ان تظهر في الابحاث الاحصائية هي الاخطاء في اجابات المستجوبين . ويمكن ان تكون هذه الاخطاء متعمدة منهم ، الامر الذي يجعل البحث يفقد قيمته فقداناً تاماً وبذلك يتحتم على الباحث اما أن يحاول تصحيح هذه الاخطاء بطريقة ما ويمكن أن يتم ذلك بأخذ عينة من المستجوبين وارسال عدادين مدربين تدريباً عالياً للاتصال بهم وتنويع الاسئلة في الاستمارة بحيث يمكن الكشف عن أخطاء المستجوب عند اجابته على سؤال معين ومراجعته بعد تنبيهه الى ذلك . ويحدث هذا النوع من الخطأ عندمـــا لا تكون هنالك الدعاية الكافية للبحث قبل النزول الى الميدان لجمع المعلومات وخاصة في المجتمعات التي ينعدم فيها الوعي الاحصائي ، الامر الذي يجعل المستجوبين يسيئون الظن بالأسئلة التي توجه اليهم فيتعمدون

الاجابة الخاطئة (السؤال عن دخل الأسرة مثلاً) ، كما ان الخوف من الحسد في بعض الدراسات يدفع المستجوبين أيضاً الى الاجابات الخاطئة (السؤال عن عدد أطفال الاسرة أو عدد الماشية والدواجن التي تملكها الاسرة) . وفي كلتا الحالتين تؤدي الدعاية للبحث خاصة اذا اتبعت جميع الوسائل الممكنة الى اتخاذ المستجوبين موقفاً لا يشوبه أي خوف من الاجابة الصحيحة .

كذلك يمكن أن تظهر الأخطاء في اجابات المستجوبين دون أي تعمد منهم وتكون بذلك أخطاء عشوائية سواء بالزيادة أو النقص ، الأمر الذي لا يغير كثيراً من المتوسطات التي نرغب في التوصل اليها فيما بعد باجراء التحليل الاحصائي على المعلومات التي نجمعها . على أن ذلك لا يجب أن يجعلنا نتهاون في جميع المعلومات الصحيحة حيث اننا لا نستطيع في حالة معينة ان نقرر ما اذا كانت أخطاء المستجوب أخطاء متعمدة أو عشوائية وحتى اذا اعتقدانها سوف تكون عشوائية لا نستطيع أن نقرر مداها . لذلك يتحتم علينا انقاص هذه الأخطاء إلى أدنى حد ممكن ، وحتى نستطيع أن ندرك الوسائل التي يمكن اتباعها في هذا الشأن يجب أن نتعرف على الأسباب التي يمكن أن تؤدي الى هذا النوع من الأخطاء . هناك أسباب كثيرة يمكن أن تؤدي الى ذلك ، منها سوء تصميم استمارة البحث بحيث تأتي الاسئلة في صيغ لا يفهمها المستجوب فهماً كاملاً صحيحاً ، كما ان لشخصية العدادين الذين يقومون بالاتصال بالمستجوبين ولتدريبهم على العمل اثر كبير على الاجابات التي يمكن التوصل اليها . كذلك يمكن أن يكون عدم شمول الحصر لوحدات المجتمع موضوع البحث قبل النزول الى الميدان لجمع المعلومات سبباً في أن تكون الاجابات التي تحصل عليها غير شاملة لجميع وحدات المجتمع ، الأمر الذي يؤدي الى تسرب الخطأ الى النتائج خاصة اذا كانت الوحدات التي لم يشملها الحصر ذات صفة مميزة لها علاقة بموضوع البحث . لذلك يكون للتصميم الجيد للاستمارة ، ولانتقاء العدادين الاكفاء وتدريبهم التدريب الكافي

وللاشراف عليهم اثنـــاء قيامهم بالعمل لمساعدتهم في مواجهة المشاكل التي يحتمل ان تصادفهم في هذا الوقت ، ولحصر وحدات المجتمع حصراً شاملاً اثر كبير على كمية ونوعية الاجابات التي نحصل عليها .

ومن ناحية اخرى يحتمل ان يجيب المستجوب اجابة صحيحة ولكن العداد نفسه يسجلها خطأ في استمارة البحث ، ويحدث ذلك اما بسبب سرعة العمل او عدم تدريب العدادين التدريب الكافي على كيفية كتابة الاجابات في الاستمارة أو عدم الاشراف الدقيق عليهم الأمر الذي يجعلهم لا يهتمون بالاتصال بالمستجوبين للحصول على اجاباتهم وانما يكتبون هم انفسهم الاجابات التي يتخيلونها رغبة منهم في عدم تحمل مشاق العمل في الميدان. كذلك يمكن أن يكون السبب في هذا النوع من الأخطاء عدم تصميم الاستمارة التصميم بحيث لا يكون فيها الفراغات الكافية لكتابة الاجابات ، الامر الذي يجعل العدادين يعملون على اختصار الاجابات المكتوبة اختصاراً مخلا فلا يكون لما هو مكتوب في الاستمارة اية قيمة عند اجراء عمليات التبويب ، ولا شك ان تدريب العدادين وشعورهم بالمسئولية ودفع الاجور المجزية لهم وتأكدهم من مراجعة المشرفين للاستمارات التي ينتهون منها وعدم قبضهم أي أجر عن الاستمارات غير الصحيحة، ومعاقبتهم بالطرد من العمل اذا ظهر ان الاجابات التي كتبوها في الاستمارات مزيفة ، كل هذه الوسائل يمكن بواسطتها انقاص هذا النوع من الاخطاء الى أدنى حد ممكن .

وعندما يكون البحث بالمعاينة يمكن أن تظهر أخطاء التسجيل بشكل خطير بحيث يفقد البحث قيمته من ناحية تحيز النتائج التي يمكن التوصل اليها. ذلك انه في مثل هذه البحوث تسحب بعض وحدات المجتمع بطريقة معينة بحيث لا يكون هناك أي تحيز في اختيارها حتى يمكن أن تكون مع بعضها صورة صادقة للمجتمع الذي تمثله . وعندما ينزل العدادون الى الميدان تعطي لهم عناوين هذه الوحدات ويطلب منهم الاتصال بها لجمع المعلومات المطلوبة ،

ولكن يحدث أحياناً ان لا يوجد المستجوب في العنوان المحدد لأسباب مختلفة منها تغيير عنوانه أو وجوده خارج مسكنه وقت زيارة العداد ، وكذلك يحدث أحياناً أن يرفض المستوجب الاجابة رفضاً باتاً ، فاذا لم تكن هناك التعليمات المشددة التي تطالب العداوين بعدم التصرف في مثل هذه الأحوال الا بعد استشارة الجهاز القائم بالبحث واذا لم يكن هناك الاشراف والمراجعة الدقيقة فان العدادين جهلاً منهم بما سوف يترتب على تصرفهم من أخطاء فادحة يستعيضون عن هذه الوحدات بوحدات اخرى يختارونها هم انفسهم دون أي اهتمام . ولا شك ان هذه الوحدات غالباً ما تكون من نوع معين ، (مثلاً اسر مكونة من رجل أعزب يسكن وحده أو رجل وزوجته والاثنان يعملان) الأمر الذي يجعل هذا النوع غير ممثل في العينة وبذلك لا تكون نتائج البحث مصورة تصويراً دقيقاً لخصائص المجتمع موضوع الدراسة . ولا يمكن معالجة هذا النوع من الاخطاء الا بتنبية العدادين عند تدريبهم الى معنى العينة وأهمية تمثيلها للمجتمع الذي تسحب منه وخطورةما يترتب على تصرفهم السابق من هذة الناحية ، هذا بالاضافة الى التعليمات المشددة التي توجه اليهم ، وحزم الجهاز القائم بالبحث في تطبيق العقوبات المنصوص عليها في هذه التعليمات ، والاشراف على العدادين أثناء عملهم في الميدان واخيراً المراجعة الدقيقة لما يقومون به من عمل .

والنوع الثالث من الأخطاء هي الاخطاء التي يمكن أن تظهر أثناء عمليات التبويب خاصة عند استعمال الآلات الاحصائية . ان استخدام الآلات الاحصائية في التبويب يتطلب عمليات مختلفة مثل الترميز والتثقيب والفرز والتبويب ، وفي كل عملية منها يمكن أن يحدث خطأ معين ولا سبيل الى انقاص هذه الأخطاء الا بمراجعة كل عملية مراجعة دقيقة واستبعاد الموظفين الذين تتكرر الأخطاء فيما يقومون به من عمل وبذلك يقتصر الجهاز على استخدام الموظفين الأكفاء .

هذه الأنواع الثلاثة من الأخطاء يمكن أن نسميها بالأخطاء العامة لأنها تظهر في كل دراسة احصائية سواء كانت بالعد الشامل أو بالمعاينة ، والصفة

الأساسية لهذه الأخطاء انه لا يمكن القضاء عليها تماماً حيث ان ذلك يعني ان كل مستجوب قد أعطى الاجابة الصحيحة الدقيقة وكل عداد قد قام بعمله على أكمل وجه ، وكل موظف في المكتب يعمل بدقة كاملة . ان البحوث الاحصائية هي من النوع الذي لا يمكن ان يقوم بها شخص واحد وحده فهي عمل مجموعة من الناس ، كما انها في الغالب تتطلب الاحتكاك بالجمهور ، ولذلك لا يمكن أن نتصور قيام مجموعة كبيرة من الناس (خاصة في التعدادات) بالأعمال التي توكل اليهم بدقة كاملة. ولذلك فان القيام بالبحث على نطاق ضيق (بالمعاينة) يساعد على انقاص هذه الأخطاء الى أدنى حد ممكن وبذلك تزداد دقة النتائج . واذا كانت هذه الأخطاء لا يمكن القضاء عليها تماماً فان انقاصها الى أدنى حد ممكن أمر يجب أن يكون هو الهدف الأول للجهاز القائم بالبحث، وكما سبق أن ذكرنا يكون التنظيم المحكم هو السبيل الوحيد الى ذلك.

على انه يجب ان نلاحظ ان اخطاء الاجابة يمكن التخلص منها تماماً اذا أمكن الحصول على المعلومات المطلوبة بطريقة موضوعية، وفي بعض البحوث يمكن التوصل الى ذلك عن طريق الحصول على المعلومات بالقياس الفعلي . فاذا كنا نقوم ببحث عن أطوال وأوزان الطلبة نستطيع عدم توجيه السؤال اليهم . ونقوم فعلاً بالقياس لهاتين الظاهرتين . كذلك اذا كنا نقوم ببحث عن المساحة المزروعة قمحاً نستطيع عدم توجيه السؤال الى حائزي الأرض عن هذا البيان ونقوم بدلاً من ذلك بقياس المساحة . بهذه الطريقة نتخلص من أخطاء الاجابة الا اذا كان هناك خطأ في المقياس الذي نستخدمه أو اذا كان الشخص القائم بالقياس غير مدرب تدريباً كافياً على استخدام اجهزة القياس التي يستعملها .

خطوات جمع المعلومات :

أشرت الى الاخطاء السابقة قبل أن أناقش خطوات جمع المعلومات حتى

يتبين لنا أهمية تنظيم هذه الخطوات في شكل متناسق متصل ، بل ان كثير من هذه الخطوات هو في الواقع نتيجة حتمية لرغبة الجهاز القائم بالبحث في انقاص الأخطاء العامة الى أدنى حد ممكن

ان أول ما يجب أن نلاحظه ان جمع المعلوما تـلا يمكن ان يتم على نطاق واسع (التعدادات) بدون أساس قانوني يركز عليه . بل انه في بعض الدول لا يمكن لاية ادارة حكومية أو غير حكومية أن تقوم بعمل احصائي فيه احتكاك بالجمهور إلا إذا كان القانون يسمح لها بذلك . لهذا يوجد في معظم دول العالم في الوقت الحاضر تشريع احصائي ينص على الآتي :

١ – حق مديرية الاحصاء وغيرها من الأجهزة الاحصائية الحكومية في جمع المعلومات . وفي بعض الدول لا بنص القانون على هذا الحق العام وانما لابد أن يصدر مرسوم خاص يحدد المعلومات التي سوف تقوم بجمعها مديرية الاحصاء وبناء على ذلك يحدد المستجوبين الذين تقع عليهم مسؤلية اعطاء المعلومات . ان النص القانوني على هذا الحق يقوى من مركز العداد ازاء المستجوب حيث لا يظهر أمامه مستطلعاً لاخباره وأسراره بل يقوم بعمله بناء على حق قانوني. وفي ذلك بعض الحافر للمستجوب على الاجابة على الأسئلة التي توجه اليه .

٢ – واجب المستجوب في كل دراسة أن يعطي المعلومات التي تطلب منه بدقة كاملة ، وعقابه إذا رفض الاجابة أو إذا تبين انه اعطى اجابات خاطئة متعمدا . فاذا كان النص الأول يحث المستجوب على الاجابة فان هذا النص يلزمه بها وبذلك لا يكون هناك من مفر أمامه الا ان يعطي المعلومات التي تطلب منه . ولا يجب أن نفهم من ذلك ان وجود هذه النصوص في التشريع الاحصائي سوف يؤدي الى اجابة جميع المستجوبين على جميع الأسئلة

التي توجه اليهم ، ذلك لأن مديرية الاحصاء لا يمكن أن تلجأ الى تطبيق هذه النصوص في كل حالة رفض إذا انها لو فعلت ذلك سوف تظهر بمظهر الذي يقاضي أفراد الجمهور وبذلك تفقد تعاونهم المخلص الأمر الذي يدفعهم الى تضليلها بكل الوسائل التي يستطيعونها .

٣ – الزام مديرية الاحصاء بالمحافظة على سرية المعلومات التي تتجمع لديها وعقاب كل موظف يثبت افشائه لسرية هذه المعلومات .

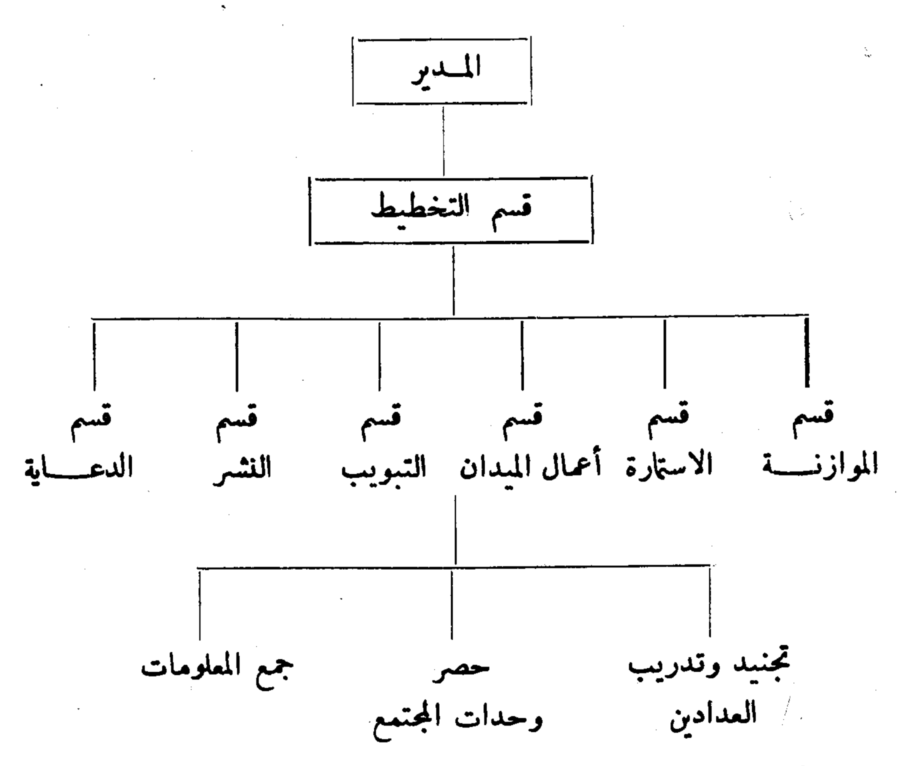
٤ – عدم استخدام المعلومات الاحصائية بشكل يعود بالضرر على صالح مستجوب معين أو مستجوبين معينين ، وبمعنى آخر لا تكون مديرية الاحصاء ملزمة بتقديم معلومات معينة تصل الى علمها عن مستجوب معين أو مستجوبين معينين الى أية جهة حكومية اخرى .

ويعني ذلك أن لمديرية الاحصاء الحق في رفض تزويد أية إدارة حكومية اخرى بمعلومات معينة عن مستجوب معين أو مستجوبين معينين ، حيث أن واجب هذه المديرية هو تقديم بيانات احصائية مجمعة لا تظهر فيها أي معالم شخصية بتاتاً . لذلك اذا وجدت مؤسسة واحدة تعمل في نشاط صناعي معين فان القانون في بعض الدول ينص على عدم نشر أية معلومات عنها الا بموافقتها الكتابية على ذلك ، فاذا رفضت تلتزم مديرية الاحصاء بعدم النشر . ويسري هذا النص كذلك في حالة وجود مؤسستين فقط تعملان في نشاط صناعي معين حيث أن نشر بياناتها وان كان في شكل ارقام مجمعة فيه افشاء لسرية بيانات كل منها .

والهدف الاساسي من هذه النصوص هو طمأنة المستجوب وحفزه على اعطاء المعلومات التي تطلب منه صحيحة دون أي تضليل ، حيث لا معنى بتاتاً لانفاق الاموال على مديرية الاحصاء اذا كانت سوف تقوم بجمع معلومات خاطئة بسبب خوف المستجوبين وعدم اطمئنانهم .

والأساس الثاني الذي يجب أن يتوفر حتى يمكن القيام ببحث ميدان هو وجود جهاز من الموظفين المدربين على العمليات المختلفة التي يتطلبها مثل هذا البحث والمتفرغين لعملهم فقط . وفي بعض الدراسات التي تكون على نطاق واسع جواً (التعدادات) يحتاج العمل الى آلاف من الموظفين ولذا لا يمكن توظيفهم جميعاً على أساس دائم حيث ان ذلك يكلف الدولة أموالاً طائلة ولهذا يكون استخدام غالبيتهم علىأساس مؤقت لمدد محدودة. والمشكلة التي يواجهها الجهاز القائم بالعمل في هذا الصدد هي في كيفية العثور على هذا العدد الكبير من الموظفين على أساس المستوى العلمي الذي يتحدد كشرط لتوظيفهم، ولذلك يتقرر في الغالب اجراء عمليات الميدان في موسم الاجازات المدرسية حيث يكون هناك عدد كبير من الطلبة والمدرسين المتفرغين والذين بذلك يمكن الاعتماد عليهم. فاذا تكون الجهاز اللازم للقيام بالعمل لا بد من تقسيم الموظفين الذين سوف يوكل اليهم الاشراف على العمليات المختلفة الى مجموعات تتخصص كل مجموعة منها في عملية معينة بشرط أن لا تستقل برأيها استقلالاً تاماً حيث ان العمليات المختلفة متداخلة ومتفاعلة مع بعضها بحيث لا بد من تحقيق التعاون بين الاقسام المختلفة التي يتكون منها الجهاز. وبالطبع يجب أن تعمل هذه الاقسام وفقاً لما يرسمه قسم التخطيط الذي يكون مسؤولاً مع مدير الجهاز عن اتخاذ القرارات الأساسية التي يحتاجها العمل ، وعن تنظيمه والاشراف عليه بحيث يضمن انتهاء كل عمليةفي الموعد المحدد لها وبذلكيضمن انهاء البحث في أقصر وقت ممكن دون الاخلال بالدقة التي تكون الهدف الأول من تنظيم العمل كما سبق أن ذكرنا . كذلك نلاحظ أن بعض الأقسام يكون عملها مركزياً فقط بينما يكون عمل البعض الآخر مركزياً ومحلياً في نفس الوقب، ذلك انه في التعدادات الواسعة النطاق لا يمكن تجميع العدادين في مكان واحد لتدريبهم نظراً لضخامة عددهم فيكون بذلك من الأفضل تجميعهم على أساس محلي. كما أن الاشراف عليهم في الميدان يكون أكثر فعالية عندما تكون صلة العدادين بالمشرفين عليهم صلة مباشرةفي المنطقة التي يعملون

فيها . والرسم التخطيطي التالي يوضح الاقسام المختلفة التي يتكون منها الجها الذي يجب تكوينه للقيام بتعداد على نطاق واسع .



والأساس الثالث الذي يجب توفره حتى يمكن ان يقوم الجهاز بعمله على الوجه الأكمل هو تخصيص الاموال التي يتطلبها العمل واعطاء مدير الجهاز سلطة انفاقها دون التقيد بالروتين الحكومي ولكن في حدود الانظمة الخاصة بالصرف . ان عدم تخصيص الاموال الكافية والتقيد بالروتين الخاص بالانفاق يعطل الجهاز في عمله ويجعله تحت رحمة المسئولين ، الأمر الذي يمكن أن يؤدي إلى فشل العملية بكاملها حيث ان العمل يجري تبعاً لجدول زمني وعدم التقيد به بسبب عجز الاموال أو بسبب الروتين يؤدي الى ارتباكه . وفضلاً عن ذلك تكون مديرية الاحصاء قد ارتبطت بموعد معين لاجراء عملية العد أمام

جمهور المستجوبين ، وارتباك العمل يؤدي الى عدم امكانية التقيد بهذا الموعد، الأمر الذي يؤدي إلى فقدان الثقة في اعمالهـــا وتدهور هيبتها . ويرتبط بتخصيص الأموال اللازمة تزويد الجهاز القائم بالعمل بالآلات الاحصائية التي بدونها لا يمكن انجاز عمليات التبويب في الوقت المناسب خاصة في التعددات الواسعة النطاق .

فاذا توفرت هـــذه الأسس الثلاث يستطيع الجهاز القيام بعمله متتبعاً الخطوات الآتية ، على انه يجب أن نلاحظ أن تسلسل هذه الخطوات لا يعنى تسلسلها فعلاً في الواقع العملي حيث سبق أن ذكرنا أن العمل يقوم على أساس توزيعه بين مجموعات من الموظفين يمكن أن يقوم كل مجموعة منها بالعمل الذي يوكل اليها في نفس الوقت الذي تقوم فيه المجموعات الأخرى بعملها . وفضلاً عن ذلك فان طبيعة العمل تستدعي أحياناً اتخاذ قرارات بشأن عدة عمليات مجتمعة في وقت واحد .

اولاً : تصميم اطار البحث وخطته العامة :

وبستدعى ذلك القيام بعدد من العمليات نوردها فيا يلي : —

١ — تحديد الهدف او الاهداف من البحث : ويعني ذلك تحديد النقاط المختلفة التي سوف يدور حولها البحث . وتحديد أهداف البحث ليس عملاً احصائياً بحتاً فلا بد أن يكون لدى الباحث ادراك واضح وفهم سليم لمدى الترابط بين هذه الأهداف ، اي لا بد أن يكون ملماً المالماً علمياً صحيحاً بموضوع البحث .

ان تحديد أهداف البحث يفيد أولاً في تحديد المعلومات المختلفة التي سوف يسأل عنها ، فلو كان أحد أهدافنـــا من بحث عن الصناعة ، مثلا ، تحديد القيمة الصافية المضافة في هذا النشاط ، فان استارة البحث لا بد وان تشتمل

على الأسئلة المختلفة التي يمكن بواسطتها التوصل إلى حساب هذه القيمة ، أما إذا كان تحديد هذه القيمة ليس من أهدافنا فلا داعي لوجود مثل هذه الأسئلة . ويجب أن نتذكر دائماً ان استمارة البحث كلما طالت واشتملت على اسئلة كثيرة كلما أدى ذلك إلى صعوبات كثيرة عند العمل في الميدان وفي المكتب (التبويب) وكلما أدى ذلك أيضاً الى مضايقة المستجوبين ، الأمر الذي قد يدفعهم الى رفض الاجابة أو التخلص منها باعطاء اجابات سريعة غير حقيقية . لذلك يجب ان يعمل الباحث كلما كان ذلك ممكناً على الا تزيد اسئلة الاستمارة كثيراً ، وتحديد أهدافه من بادىء الأمر يساعده في عدم توجيه أسئلة لا يكون البحث في حاجة اليها .

كذلك يساعد تحديد أهداف البحث في تحديد المجتمع الاحصائي موضوع البحث أي في تحديد مجموع الوحدات التي سوف تعتبر وحدات استبيان عند النزول إلى الميدان لجمع المعلومات . فإذا كنا نريد القيام ببحث عن ميزانية الأسرة ، مثلاً ، وكان هدفنا من البحث استنتاج الاوزان التي ترجح بها أسعار السلع التي تدخل في تركيب الرقم القياسي لنفقة معيشة الطبقة العاملة (سنناقش ذلك بتفصيل في موضوع الارقام القياسية) فان مجتمع البحث يتحدد تبعاً لذلك بأسر الطبقة العاملة التي لا يزيد دخلها الشهري عن مبلغ معين ، أما إذا كان هدفنا من البحث التوصل الى تقدير استهلاك المجتمع عامة من السلع المختلفة فان مجتمع البحث في هذه الحالة لا بد أن يشتمل الاسر عامة بصرف النظر عن مستوى دخلها . كذلك في الحالة الاولى فان أسئلتنا سوف توجه إلى قيم الانفاق على السلع المختلفة ، أما في الحالة الثانية سوف تكون الاسئلة عن الكميات المستهلكة من السلع المختلفة في مدة معينة .

وفي الابحاث التجريبية تتخذ هذه الخطوة شكلاً آخر حيث يتعين على الباحث تحديد وتعريف المشكلة التي سوف يدور حولها البحث . وفي هذه

الحالة أيضاً يؤدي تحديد وتعريف المشكلة الى تحديد المعلومات التي سوف يسأل عنها والمجتمع الاحصائي الذي سوف تجرى عليه التجربة .

وقبل تحديد المعلومات التي يريد الباحث ان يسأل عنها يجب دراسة كل المراجع الممكنة للوقوف على مـــا جمع من هذه المعلومات فعلاً في دراسات سابقة ، حيث ان ذلك يساعد في اختصار استمارة البحث وهذا يوفر كثيراً من الجهود والتكاليف .

٢ – تحديد نطاق وشمول البحث ، أي تحديد وتعريف المجتمع موضوع البحث ويساعد في ذلك كما قدمنا تحديد الأهداف . على انه لا يكفي تحديد المجتمع بشكل عام بل يجب تعريفه تعريفاً دقيقاً ويتم ذلك بتحديد وتعريف الوحدة التي يتكون منها هذا المجتمع . ففي بحث عن الصناعة يكون المجتمع الاحصائي هو مجموع المؤسسات الصناعية في الدولة ، ولكن ما هو تعريف المؤسسة الصناعية الذي يحدد لنا هذا المجتمع تحديداً دقيقاً ، هل هي المؤسسة التي تقوم بالانتاج الصناعي أو الاصلاح أو الصيانة في مكان محدد وقت جمع المعلومات ، أم هي المؤسسة التي تقوم بالانتاج الصناعي فقط وبذلك لا يدخل في المجتمع الذي نبحثه المؤسسات التي يقتصر نشاطها على الاصلاح والصيانة فقط . واذا كنا نقوم ببحث عن الزراعة يكون المجتمع الاحصائي هو مجموع الحيازات الزراعية في الدولة ، ولكن هل يدخـــل في هذا المجتمع مجموع الحيازات بشكل عـام مهما كانت صغيرة أو كبيرة أو سنقتصر فقط على الحيازات التي تزيد مساحتها عن مساحة معينة . وكذلك في دراسة عن السكان هل يدخل في تعريفنا لمجتمع البحث السكان الرحل أيضاً ، أم سوف نستبعدهم.

ومن الواضح أن وضع تعريف معين لوحدة المجتمع الاحصائى الذي نقرر دراسته لا يعني اننا سوف نجد ان هـــذا التعريف ينطبق في الواقع العملي انطباقاً كاملاً ، حيث انه في الغالب نصادف مشاكل معينة عند ربط التعريف

النظري الذي نستخدمه بالواقع العملي . ومهمة الباحث هي توقع هذه المشاكل وايجاد حل لها حتى يكون العداد في عمله في الميدان على بينةمن أمره . كما أن الباحث يجب أن ينبه العدادين الى انهم قد يصادفون مشاكل اخرى غير تلك التي نبههم اليهم ، فاذا حدث ذلك لا يجب أن يتصرفوا الا بعد أخذ رأي القائمين على الاشراف عليهم ، وبهذه الطريقة يمكن حل المشاكل المختلفة التي تصادف العدادين بالنسبة لتعريف وحدة الاستبيان حلاً موحداً . فاذا كان بحثنا عـن الصناعة واستخدمنا التعريف الذي سبق الاشارة اليه للمؤسسة الصناعية فسوف نجدفي الواقع العملي مؤسسات تعمل في انشطة اقتصاديةمختلفة، مثلاً مؤسسة تربي الماشية وتأخذ منها الحليب فتصنعه ثم تبيعه الى المستهلكين، هذه المؤسسة تعمل في النشاط الزراعي والصناعي والتجاري ، هل تدخل في المجتمع الاحصائي موضوع البحث أو تستبعد ؟ واذا ادخلناها كيف يمكن جمع المعلومات الخاصة بنشاطها الصناعي فقط ؟ هذه الأسئلة وغيرها تعبر عن المشاكل العملية التي تظهر عند تطبيق تعريف وحدة الاستبيان وواجب الباحث أن يضع لها الحلول حتى يكون عمل العدادين جميعـــاً على أساس موحد .

٣ – هل يكون البحث آنياً أو على مراحل ؟ والمقصود بالبحث الآني توحيد الفترة الزمنية التي تجمع فيها المعلومـــات من جميع وحدات المجتمع الإحصائى الذي تقرر دراسته . ومن الواضح ان إجراء العمل بهذه الطريقة يستدعي استخدام عدداً كبيراً جداً من العدادين ، الأمر الذي يوقع الجهاز القائم بالعمل في مشاكل كثيرة ، منها الاضطرار إلى استخدام عدادين من مستوى علمي متدني ، وصعوبة تدريبهم التدريب الكافي ، وصعوبة الاشراف عليهم في الميدان ، وصعوبة تنظيم العمل الميداني نفسه في الفترة المحددة لجمع المعلومات ، ثم استقبال المكتب لعدد كبير جداً من الاستمارات بعد الانتهاء منها في الميدان في وقت واحد ، الأمر الذي يؤدي الى ارتباك العمليات

المكتبية المختلفة (المراجعة والترميز الخ) . ولا شك ان الجهاز يحاول معالجة هذه الصعوبات بما يراه مناسباً ، الا أنه مهما كان ماهراً في ذلك لا بد وان يترتب على العمل بهذه الطريقة أخطاء كثيرة تنقص من دقة النتائج الأخيرة . وتتبع هذه الطريقة عندما يكون المجتمع الاحصائي موضوع البحث مجتمعاً متحركاً أي عندما تكون معلوماته قابلة للتغيرات السريعـة مثل المجتمع السكاني ، فالسكان في حركة دائبة تشمل مواليد جدد ووفيات وتغيرات مستمرة في الحالة الزواجية والتعليمية ... الخ ولهذا يضطر الجهاز القائم بتعداد السكان إلى أن يجري العمل بهذه الطريقة بالرغم من معرفته للصعوبات المختلفة التي سوف تواجهه معتمداً على الوسائل المختلفة التي يحاول اتباعها للتخفيف من حدة هذه الصعوبات ولانقاص الأخطاء التي يمكن ان تترتب عليها .

أما إذا كان المجتمع موضوع البحث لا يتعرض لهذه الحركة الدائبة يكون من الأفضل اجراء العمل على مراحل . وذلك بتقسيم المجتمع إلى أقسام والبدء بجمع المعلومات من قسم معين أولا ثم ينتقل فريق العدادين المكلفين بالعمل الى الأقسام الاخرى تباعـاً . ان العمل بهذه الطريقة يقضي على كثير من الصعوبات التي سبق الاشارة اليها ، حيث يكون العمل في حاجة الى عدد قليل من العدادين يمكن انتقائهم على أساس المستوى العلمي اللازم للبحث وعلى أساس شخصية العداد ولباقته وبذلك يمكن التثبت من عنصر هام من عناصر نجاح العمل ، كذلك يكون تنظيم العمل الميداني والاشراف عليه أسهل بكثير من البحث الآني ، كما ان الاستمارات التي تنتهي في الميدان ترد إلى المكتب تباعاً وعلى مراحل الأمر الذي يخفف من ضغطها ، وبالاضافة الى ذلك يزداد العدادون خبرة بعملهم كلما انتقلوا من قسم من المجتمع إلى قسم آخر .

ويرتبط بهذا القرار كذلك تحديد موعد النزول إلى الميدان لجمع المعلومات وهل تجمع المعلومات مرة واحدة أو على فترات دورية . هناك عوامل كثيرة

يجب أن تؤخـــذ في الاعتبار عند تحديد موعد النزول إلى الميدات لجمع المعلومات ، منها مدى الاحتياج إلى العدادين المتفرغين فكلما زاد الاحتياج كلما أصبح من الواجب أن يكون موعد جمع المعلومات خلال العطلة المدرسية الصيفية . كما ان ظروف الطقس يمكن أن يكون لهـــا تأثيرها خاصة في الدراسات الزراعيه إذ لا يجب النزول إلى الميدان في موسم الامطار وبشكل خاص عندما تكون طرق المواصلات في المناطق الزراعية في حالة رديئة ، كذلك لا يجب النزول إلى الميدان في الدراسات السكانية في المواسم التي يكون السكان فيها في حركة شاذة بسبب الاصطياف مثلا . وفي البحوث الخـاصة بميزانيه الاسرة يجب جمع المعلومات على فترات دورية حتى يمكن جمع معلومات تمثل التغيرات الموسمية في انفاق الاسر . كذلك في البحوث التي يكون الهدف منها معرفة التغيرات الجارية في الظواهر موضوع البحث مثل التوظف والبطالة والاجـــور والانتاج الصناعي يكون من الواجب جمع المعلومات في فترات دورية .

٤ — تحديد طريقة جمع المعلومـــات — وتجمع المعلومات عادة باحدى طريقتين أساسيتين ، هما طريقة الادلاء الشفهي وطريقة التسجيل الذاتي . وفي الطريقة الاولى يقوم العدادون بالمقابلة الشخصية للمستجوبين ويوجهون اليهم الأسئلة ثم يسجلون إجاباتهم في استمارة البحث . والاستمارة في هذه الطريقة تسمى كشف البحث . أما الطريقة الثانية فتقوم على أساس الاتصال غـير المباشر بالمستجوبين وذلك بواسطة البريد أو التليفون. والاستمارة التي تستخدم في هذه الطريقة تسمى صحيفة الاستبيان .

على ان العمل يجري في الغالب بالجمع بين الطريقتين حيث يقوم العدادون بتوزيع الاستمارات على المستجوبين ويطالبونهم بملأها إذا تبينوا منهم القدرة على القراءة والكتابة إلى الدرجة التي تمكنهم من القيام بذلك ، ثم يزورونهم مرة ثانية لجمـــع الاستمارات ومساعدة المستجوبين في اجابة الأسئلة التي لم

يتمكنوا من فهمها أو التي لم يتمكنوا من فهم كيفية كتابة الاجابة عليها . كذلك يمكن أن ترسل الاستمارات إلى المستجوبين بالبريد وبعد ورود الاستمارات الى المكتب يمكن تحديد المستجوبين الذين لم يهتموا باعادة استماراتهم فيقوم العدادون بزيارة هؤلاء المستجوبين للحصول على إجاباتهم ، وبذلك نستطيع أن نقول أن الجهاز القائم بالعمل لا يجب أن يتقيد بطريقة معينة تقيداً جامدا ، إذ يجب أن يتبع ما تمليه عليه طبيعة البحث وتطوره .

وبالنسبة لبعض البحوث تكون الطريقة التي يجب اتباعها هي الملاحظة حيث يقوم العدادون بملاحظة الظاهرة موضوع البحث ثم كتابة المعلومات دون توجيه الأسئلة كما هو الحال في دراسة عن كيفية استجابة الاطفال في سن معينة مثلاً لأنواع معينة من المؤثرات ، وكذلك في دراسة عن حركة السير حيث يقوم العدادون بملاحظة السيارات ثم تدوين ملاحظاتهم في الاستمارة المعدة لذلك . وبالنسبة لبعض البحوث الاخرى تكون الطريقة التي يجب اتباعها هي التسجيل وذلك في الحالات التي تكون فيها الهيئة المشرفة على البحث لها من السلطة التي تعطيها الحق في الزام المبحوثين بتقديم البيانات المطلوبة لتسجيلها في سجلات معينة (سجلات الجمارك — سجلات المواليد — سجلات الموظفين .. الخ) .

ولكل من كشف البحث وصحيفة الاستبيان مزاياه وعيوبه ، وتعتبر مزايا كل منهما عيوباً للآخر والعكس . ومن أهم مزايا كشف البحث ما يلي :

١ — يساعد كشف البحث على جمع بيانات بالمشاهدة المباشرة لبعض الظواهر دون الحاجة الى توجيه اسئلة عنهاخاصة اذا كانت هذه الاسئلة منفرة أو محرجة ولا يحتمل الحصول على اجابات دقيقة عليها فمثلاً يستطيع العداد عند زيارته لاسرة ما ان يشاهد حالة أثاث المنزل ووسائل الاضاءة وملابس الأفراد وغير ذلك من البيانات التي لا يمكن السؤال عنها لما قد يكون في ذلك

من أثر سيء في نفس المستوجب قد يجعله يرفض الاجابة على باقي الاسئلة أو يعطي اجابات غير دقيقة .

٢ – تصلح هذه الوسيلة في الدراسات التي تتطلب اسئلة كثيرة العدد، اذ ان العداد يستطيع بلباقته أن يكسب معونة المستجوبين باقناعهم بما لهذه الدراسة من نتائج تعود على المجتمع بالنفع والفائدة ، كما يستطيع أن يسهل مهمتهم بشرح الاسئلة بما يتناسب وادراك كل منهم وصرف النظر عن الاسئلة التي يبدي المستجوب عدم الرغبة في الاجابة عليها .

٣ – ليس من المستحب ولا من المتيسر أن يكلف المستجوبين بقراءة تعليمات مطولة عن معاني الاسئلة وكيفية الاجابة عليها اذ قد لا يكون لديهم الوقت الكافي او الاستعداد لقراءة هذه التعليمات . لهذا يعتبر كشف البحث من انجح الوسائل لجمع المعلومات في حالة الدراسات التي تتطلب توجيه اسئلة يكون من الصعب الاجابة عليها دون شرح خاص حيث ان العداد يمكن أن يقوم بتوضيح معاني الكلمات وشرح المقصود من الاسئلة كلما استدعى الامر ذلك.

٤ – من الملاحظ أن الشخص لا يجيب على السؤال الذي يوجه اليه إلا تبعاً للمعنى الذي يفهمه من هذا السؤال . فاذا فهم معنى يخالف المقصود فعلاً من السؤال كانت اجابته عديمة القيمة بل قد تؤدي الى نتائج مضللة وغير دقيقة. ومن النادر ان يحدث ذلك في حالة كشف البحث ، اذ انه اذا اتضح للعداد ان المستجوب أعطى اجابة غير متفقة مع المعنى المقصود من السؤال فانه يقوم بشرحه وتوضيحه حتى يحصل على الاجابة الدقيقة له .

٥ – اذا شك العداد في صحة الاجابات التي يحصل عليها أو إذا اتضح له تناقض الاجابات تناقضاً ظاهراً فانه يستطيع مناقشة المستجوب حتى يحصل على بيانات متماسكة لا تناقض فيها .

٦ – باستخدام كشف البحث كاداة لجمع المعلومات يكون من المؤكد إلى درجة بعيدة الحصول على اجابات لكل الأسئلة من جميع المستجوبين (أو على

الأقل من نسبة كبيرة منهم) خاصة إذا تحققت جميع العوامل الأخرى التي تساعد على نجاح الدراسة مثل اختيار العدادين بحيث يصلحون لهذا العمل من جميع النواحي وتدريبهم التدريب الكافي على كيفية اجراء العمل بدقة والقيام بالدعاية اللازمة بجميع الطرق الممكنة في حالة الدراسات الواسعة النطاق ، وإذا أحسن العدادون عرض أغراض الدراسة على المستجوبين واقناعهم باهمية تعاونهم ، وإذا اختير الوقت المناسب للاتصال بالمستجوبين .

إلا أن كشف البحث له عيوب أهمها : —

١ — كثرة التكاليف والمجهودات التي تبذلها الهيئة المشرفة على الدراسة في اختيار العدادين وتدريبهم ودفع مرتباتهم واجور انتقالهم .

٢ — يكون للعدادين سلطة كبيرة على الاجابات التي تؤخذ من المستجوبين وحتى يمكن أن تأتي هذه الاجابات صحيحة ودقيقة يجب أن يتصف العدادون بالحياد التام ، أما إذا كان بعض العدادين متحيزين لافكار معينة (خاصة في الدراسات الاجتماعية) فان اتصالهم الشخصي بالمستجوبين يتيح لهم فرصة التأثير عليهم وتوجيه الأسئلة اليهم بحيث تأتي الاجابات مؤيده لافكارهم . وهذا بالطبع يتنافى مع الروح العلمية التي يجب أن تتصف بها الدراسة .

٣ — في بعض الدراسات تشتمل استمارة البحث على أسئلة قد تضر بالمستجوب الذي يدلي بها ضرراً مادياً أو أدبياً . وذلك مثل الاسئلة عن تعاطي المخدرات أو العلاقات الجنسية أو اعتناق المبادىء السياسية التي تحرمها الدولة وما شابه ذلك ، وفي هذه الدراسات لا يمكن أن يصلح كشف البحث في الحصول على اجـــابات لاسئلة من هذا النوع نظراً للمقابلة الشخصية بين العداد والمستجوب .

أما صحيفة الاستبيان فلا تختلف كثيراً عن كشف البحث إلا في طريقة ملئها فهي اما أن تسلم للمستجوب أو ترسل اليه بالبريد أو تنشر على صفحات

الجرائد والمجلات (في بعض الدراسات) ويقوم المستجوب بتدوين اجاباته على اسئلتها ثم يعيدها إلى الهيئة المشرفة على الدراسة . ولصحيفة الاستبيان كما لكشف البحث مزايا وعيوب أهمها من ناحية المزايا ما يلي :

١ – قلة التكاليف والمجهودات اللازمة لجمع البيانات خاصة إذا أرسلت بالبريد أو نشرت على صفحات الجرائد حيث يمكن توفير العنـــاء في تجنيد العدادين وتدريبهم ودفع اجورهم وغير ذلك من نفقات .

٢ – تصلح في الدراسات التي تتطلب الحصول على معلومات لا يمكن الجهر بها لمـا قد ينتج من ذلك من ضرر أو لأن التقاليد لا تسمح بالتصريح بمثل هذه البيانات . حيث باستخدام صحيفة الاستبيان يمكن عدم كتابة اسم المستجوب وبذلك لا يمكن التعرف على صاحبها الأمر الذي يشجعه على اعطاء الاجابات المطلوبة بصراحة تامة .

٣ – كما انه باستخدام صحيفة الاستبيان لا تكون الدراسة تحت رحمة العدادين من ناحية اهمالهم أو تحيزهم .

أما عيوب صحيفة الاستبيان فيمكن تلخيصها في الآتي :

١ – لا يمكن استخدامها إلا إذا كان المستجوبون مثقفين أو على الأقل يجيدون القراءة والكتابة .

٢ – يكون تصميم صحيفة الاستبيان من الصعوبة بمكان كبير حيث يجب أن تكون الاسئلة في غـاية البساطة والوضوح ولا تحتمل التأويل حيث لن يكون هناك عدادون يساعدون المستجوبين في فهم الاسئلة الفهم الصحيح .

٣ – لا يمكن التوسع في الاسئلة الموجهة الى المستجوبين لأن ذلك يمكن أن يؤدي الى اهمال الاجابة عليها .

٤ – إذا كانت الدراسة تستدعي عدم كتابة اسم المستجوب لتشجيعه

على اعطاء الاجابات الصحيحة ، فانه لا يمكن الرجوع اليه لاستكمال ما قد يكون بالصحيفة من نقص أو لتصحيح الاخطاء أو الاستفسار عن الاجابات المتناقضة ان وجدت .

٥ - قلة عدد صحائف الاستبيان التي ترد إلى الهيئة المشرفة بعد ملئها مهما وضعت هذه الهيئة من مغريات وتسهيلات للاجابة خاصة إذا كانت الصحائف ستعاد بالبريد (هذا بالمقارنة مع كشف البحث) . ولذلك أهمية كبيرة خاصة إذا كانت الدراسة تجري بالمعاينة حيث يصبح عدد الاجابات التي تحصل عليها الهيئة المشرفة على الدراسة غير متفقاً مع حجم العينة الذي حدد أولاً ، كما ان الصحائف التي تعـــاد إلى الهيئة لا يمكن اعتبارها ممثلة للمجتمع موضوع الدراسة، الأمر الذي يدخل عنصر التحيز في النتائج واحياناً يتضخم هذا العنصر فتفشل الدراسة . ويكون هذا التحيز راجعاً الى اهتمام البعض بموضوع الدراسة لسبب معين ، مثلا ، محاولة استخدامالدراسة كفرصة لاظهـــار رأيهم وعدم اهتمام البعض الآخر ، الأمر الذي يدفعهم إلى عدم الاجابة . ولعل ذلك هو السبب في فشل الدراسات الخاصة بالاستقصاء عن نتائج الانتخابات في بعض الدول ، حيث يكون المؤيدين لمرشح ما واثقين من نجاحه فلا يهتمؤن بالاجابة على الاستقصاء، بينما يتخذ المعارضون الاستقصاء وسيلة للتعبير عن اتجاههم فيجيبون ، ولهذا تأتي النتيجة غير متفقة مع ما يحدث فعلا .

٦ - أما بالنسبة للصحائف التي تعاد فعلا الى الهيئة المشرفة على الدراسة فكثير منها سوف يشتمل على اجابات تهكمية أو مغرضة ، وبالطبع لن تكون لهذه الصحائف أي قيمة بل ستؤدي إلى ضياع الوقت في قراءتها .

على انه يمكن استخدام بعض الأساليب والوسائل إلى التقليل من أثر العيوب السابقه ومن أهم هذه الأساليب العمل على توضيح أغراض الدراسة توضيحاً جذاباً وبكل الوسائل الممكنة (خاصة في الدراسة الواسعة النطاق)

بحيث تظهر الفوائد العملية والعلمية لنتائجها الدقيقة ، وكذلك يمكن أن يرفق بالصحيفة عند ارسالهـــا للمستجوب خطاب رقيق يوضح أغراض الدراسة بامانة ودقة ويبين للمستجوب أهمية تعاونه باعادة الصحيفة بعد ملئها والفائدة التي تتحقق باعطـــاءه الاجابات الصحيحـــة والدقيقة . ولتسهيل مهمة المستجوب اذا طلب منه اعادة الصحيفة بالبريد يحسن أن يرسل اليه مع الصحيفة مغلف عليه طابع بريد وعنوان الهيئة المشرفة على الدراسة ، ويا حبذا لو تضمن المغلف على أكثر من نسخة من الصحيفة لمواجهة احتمال تلف احداها .

٥ – تحديد مجال جمع المعلومات ، أي هل سيشمل جمع المعلومات جميع وحدات المجتمع الاحصائي الذي تقرر بحثه (العدد الشامل) ، أو عينة من هذا المجتمع ثم تعميم نتائجها على المجتمع الأصلي ؟ ولا شك أن هناك اسباب تؤدي الى هذا السؤال، الأمر الذي يقودنا الى مناقشة مزايا الدراسة بالمعاينة.

يؤدي استخدام المعاينة الى توفير جزءاً كبيراً من التكاليف والجهود التي نضطر اليها في حالة العد الشامل حيث اننا سنستخدم جزءاً من المجتمع وليس كله . على انه يجب أن نلاحظ ان توفير التكاليف لا يكون بنسبة حجة العينة الى حجم المجتمع الذي سحبت منه، فاذا كان حجم العينة يكون ١٠٪ من حجم المجتمع فلا يترتب على ذلك انخفاض التكاليف الى العشر حيث ان الدراسة بالمعاينة تستدعي اجراء خطوات عمل وخبرة رياضية لا يحتاجها العد الشامل ، الأمر الذي يوجد عناصر جديدة من التكاليف . ولهذه الميزة أهمية كبيرة حيث أن كثيراً من البحوث لو لم يكن من الممكن اجرائهـــا بالمعاينة لما قامت بها الهيئات المعنية نظراً لارتفاع التكاليف التي لا قبل لها بتحملها اذا كان لا بد من اجرائهـــا بالعد الشامل ، ولعل ذلك يفسر كثرة البحوث الاحصائية في الوقت الحاضر وخاصة في الدول التي لا تستطيع أن تتحمل ميزانياتها نفقات العد الشامل الباهظة . كما ان لتوفير الجهود اهميـة

كبيرة حيث ان المستوى التعليمي في بعض الدول لا يسمح لها بتزويد الاجهزة الاحصائية بالعدد اللازم من الموظفين الذين تحتاجهم إذا كان يتحتم عليها اجراء جميع الدراسات الاحصائية بالعد الشامل . وبشكل عام نستطيع أن نقول أن امكانية اجراء البحوث الاحصائية بالمعاينة فتحت مجالات واسعة للبحث لم يكن من الممكن تصورها من قبل .

كذلك يؤدي استخدام المعاينة الى توفير الوقت الذي يضيع في اجراء البحوث بالعد الشامل ، ولتوفير الوقت أهمية كبيرة، ففي كثير من الدراسات يؤدي اتساع نطاق العمل بالعد الشامل الى ظهور النتائج بعد سنوات طويلة (تعداد السكان مثلا) بحيث تصبح هذه النتائج صورة لما كان عليه المجتمع قبل بضع سنوات ، أي تصبح مجرد صورة تاريخية لا يكون لها أهمية كبيرة في اتخاذ اجراءات عملية على أساسها (التخطيط الاقتصادي والاجتماعي مثلاً) . وفي بعض الحالات تكون الدولة في حاجة الى بيانات احصائية بعد وقت قصير حتى تستطيع أن تقرر سياستها نحو الظاهرة المتعلقة بهذه البيانات ، مثلاً عندما ترغب الحكومة في تقدير لمحصول القمح في عام معين حتى إذا تبينت عجزه بدأت تعمل على استيراد ما تحتاجه من الخارج والعكس إذا تبينت زيادته عما يكفي المجتمع تبدأ في محاولة تصريف الفائض منه في الأسواق الخارجية . وفي مثل هذه الحالة يتحتم اجراء البحث بالمعاينة نظراً لأن اجرائه بالعد الشامل لا يتفق مطلقاً مع الغرض منه .

كذلك يمكن أن يكون البحث أكثر تفصيلاً في الدراسات بالمعاينة ، حيث انه في حالة اجرائه بالعد الشامل يأخذ الباحث اتساع نطاق البحث في اعتباره فيحاول دائماً انقاص عدد الاسئلة في الاستمارة لأن زيادتها يزيد المشكلة تعقيداً وتكاليفاً وجهداً . ان اجراء البحث بشكل تفصيلي يحتاج في الغالب إلى عدادين اكفاء متخصصين ومدربين تدريباً عالياً على العمل الذي يوكل اليهم. ومن الواضح انه لا يمكن أبداً الحصول على هذا النوع من العدادين في الدراسات

بالعد الشامل . واذا أخذنا في اعتبارنا جميع خطوات العمل من حيث الوقت والجهد والتعقيد أدركنا انه يستحيل اجراء البحوث بشكل تفصيلي بواسطة العد الشامل . لذلك تساعدنا الدراسة بالمعاينة على اجراء بحوث أكثر تفصيلاً وأكثر ارتباطاً بالواقع تبعاً لذلك ، الأمر الذي يزيد من اهميتها العلميـــة . وقد أدركت الاجهزة الاحصائية في كثير من الدول هذه الميزة فعملت على اجراء دراسات بالمعاينة بجانب تعداد السكان الواسع النطاق وأثناء القيام به للتعرف على الصفات التفصيلية للسكان لمساعدة الحكومة في تخطيط الشئون المختلفة الخاصة بهم .

كذلك يمكن إنقاص الأخطاء العامة التي سبق الاشارة اليها إلى أدنى حد ممكن بسبب ضيق نطاق العمل . حقاً ان الدراسات بالمعاينة تؤدي الى ظهور نوع آخر من الخطأ لا يظهر في العدّ الشامل وهو الخطأ الكامن في عملية المعاينة نفسها ، حيث نلاحظ انه مهما دققنا في اختيار وحدات العينة ومهما اتبعنا في ذلك من طرق صحيحة فان العينة الناتجة لا يمكن ان تكون صورة طبق الأصل من المجتمع الذي سحبت منه ، وبمعنى آخر لا يمكن ان نقرر ان المقاييس التي نصل اليها عن طريق المعاينة تكون هي ذاتها المقاييس (المتوسط مثلا) التي كان يمكن أن نصل اليها لو أجرينا الدراسة بالعد الشامل ، فلا بد أن يوجد فرق وهو الذي نسميه خطأ المعاينة ، وينشأ هذا الخطـــأ بسبب اختلاف وزن الأنواع المختلفة من الوحدات في كل من المجتمع والعينة ، وغالباً تكون الوحدات الشاذة ذات وزن قليل في المجتمـــع نظراً لضخامة عدد وحداته وذات وزن اكبر في العينة نظراً لقلة عدد وحداتها . على ان النظرية الرياضية الخاصة بالاحتمالات تساعدنا في حساب هذا الخطأ وهو أمر من الأهمية بمكان كـــبير حيث انه اذا لم يكن من الممكن حساب هذا الخطأ لكانت الدراسة بالمعاينة عديمة الأهمية بتاتاً ، فنحن نعلم في هذه الحالة بوجود فرق بين المقاييس الناتجة من العينة وتلك التي تنتج من العد الشامل ولكننا لا نعلم كم هو هذا الفرق ، الأمر الذي يجعل دراستنا عمياء لا قيمة لها .

وبذلك إذا افترضنا مثلاً ان الأخطاء العامــة في العد الشامل تصل الى ١٠ ٪ ، وأمكن إنقاصها بالمعاينة الى ٢ ٪ ، وإذا افترضنا ان خطأ المعاينةهو ١ ٪ يكون مجموع الخطأ في نتائج المعاينة هو ٣ ٪ فقط بالمقارنة مع أخطاء العد الشامل التي تصل الى ١٠ ٪ ؛ لذلـك تكون الميزة الاساسية للأبحاث بالمعاينة هي زيادة دقة النتائج خاصة في الدراسات التي تدور حول موضوعات معقدة ، إذ انه في هذه الدراسات يكون هناك مجال واسع لانقاص الأخطاء العامة حيث نستطيع أن نستفيد من ارتفاع مستوى العدادين ومن اعطائهم الوقت الكافي لشرح كل سؤال شرحاً جيداً للمستجوبين .

ولقد لاحظنا مثلاً أن أخطاء الاجابة يمكن القضاء عليها اذا امكن جمع بيانات موضوعية بالقياس الفعلي للظواهر موضوع البحث (قياس مساحات الأراضي الزراعية مثلاً) ومن الواضح انه لا يمكن استخدام هذه الطريقة في حالة الدراسات بالعد الشامل حيث لا يكون هناك الوقت الكافي لذلك .

كذلك هناك مجتمعات يكون من المستحيل بأي حال من الأحوال اجراء بحثها احصائياً بالعــد الشامل مثل الأسماك والطيور والحيوانات المفترسة الخ ، وبذلك لا يكون هناك مفراً من استخدام المعاينة ، كذلك لا يجب أن ننسى ان التجارب العلمية ما هي الا دراسة بالمعاينة .

كذلك يمكن أن يتبين لنا بعد اجراء البحث بالعد الشامل ان المستجوبين تحيزوا في اجابة سؤال معين وليس هناك وسيلة لتصحيح هذه الاجابات الا بأخذ عينة منهم ومعاودة توجيه السؤال بروية واهتمام أكبر ، ولعل في ذلك ما يساعدنا على الحصول على نتائج أكثر دقة .

لهذه الأسباب مجتمعة انتشر استخدام العينات في مختلف ميادين البحث العلمي واستعيض بها عن العد الشامل ، الأمر الذي دفع علماء الاحصاء إلى تطوير الأبحاث الخاصة بهذه الطريقة في البحث حتى أصبحت تكون فرعاً مستقلاً في علم الاحصاء . على ان المزايا الكثيرة التي تترتب على استخدام

المعاينة لا تعنى ان اتباعها لا يواجهنا بأية صعوبات ، ذلك لأن هنالك مبادىء عامة اساسية يجب ان نأخذها دائماً بعين الاعتبار عند بحث هذا الموضوع ويمكن تلخيصها في الآتي :

١ – ان سحب العينة التي يمكن أن تكون صورة صادقة للمجتمع الذي سحبت منه يستلزم معرفة بعض خصائص هذا المجتمع مقدماً حتى يمكن تحديد نوع العينة التي تناسبه وحجم العينة التي يجب سحبها والطريقة التي يجب اتباعها في عملية السحب وبدون هذه المعرفة تصبح المعاينة أمراً متعذراً مستحيلا . ان المجتمعات التي يمكن أن نرغب في بحثها احصائياً ليست جميعاً من نفس النوع، ففي بعض هذه المجتمعات نلاحظ التشابه الكبير بين وحداتها الذي يمكن أن يصل إلى درجة التجانس ، وفي بعضهـــا نلاحظ الاختلاف الكبير ، وحيث أن العينة يجب أن تكون صورة صادقة من المجتمع الذي تسحب منه فمن البديهي ان حجم العينة بالنسبة للمجتمعات التي هي من النوع الأول يمكن أن يكون حجماً صغيراً بينما لا بد أن يكون حجماً كبيراً بالنسبة للمجتمعات التي هي من النوع الثاني ، وبمعنى آخر يتناسب حجم العينة تناسباً طردياً مع درجة الاختلاف بين وحدات المجتمع الذي تمثله . وبذلك لا بد أن نتعرف على درجـــة الاختلاف بين وحدات المجتمع حتى نستطيع أن نحدد حجم العينة التي يجب سحبها منه للتوصل إلى نتائج ذات دقة معينة . هذا المثال البسيط يوضح لنا أن سحب العينات لا يمكن أن يجري مستقلا بذاته دون بعض المعرفة لخصائص المجتمعات التي نريد دراستها بالمعاينة .

٢ – ان الهدف الأساسي عند سحب العينات هو التوصل إلى مجموعة من الوحدات ، التي يمكن أن تكون مع بعضها صورة صادقة من المجتمع الذي نرغب في دراسته حتى يمكن ان تكون اخطاء المعاينة في النتائج اقل مـــا يمكن وبذلك تزداد اهميتها كمقاييس يمكن الاعتماد عليها في النواحي العملية المختلفة

(التخطيط مثلا) . لذلك يواجه الباحث عند سحب اية عينة مشكلة اساسية تتلخص في كيفية منع التحيز في الاختيار حتى يستطيع ان يطمئن الى العينة المسحوبة . لذلك يكون من المفيد ان نناقش العوامل المختلفة التي يمكن ان تودي الى التحيز في العينة حتى نستطيع ان نتفاداها عند قيامنا بالعمل .

أ - استخدام اطار [1] غير شامل لوحدات المجتمع جميعاً ، ولا شك ان التحيز الناتج يكون خطيراً اذا كانت الوحدات التي لا يشملها الاطار من نوع معين له علاقة بموضوع البحث ، مثلاً اهمال الاسر التي تسكن في العشش في دراسة عن مستوى المعيشة بين اسر مدنية ما . ان عدم ادراج هذه الاسر ضمن الاطار لا يعطيها

---

(١) اطار المعاينة هو مجموع الوحدات التي نسحب منها العينة، أي مجتمع المعاينة، الوحدة في هذا الاطار هي وحدة المعاينة . ويمكن أن يكون مجتمع المعاينة هو نفسه المجتمع الاحصائي موضوع البحث أو مجتمع آخر يختلف عنه ، فمثلا عند سحب عينة من الأسر يمكن أن يكون الاطار عبارة عن قائمة باسماء اسر المجتمع موضوع البحث وعناوين هذه الاسر، ويمكن أن يكون الاطار عبارة عن خريطة تظهر فيها بلوكات المساكن التي تسكنها هذه الاسر ، وكل بلوك منها يكون هو وحدة اللعاينة ومجموع البلوكات هو مجتمع اللعاينة . وبالطبع يسحب عدد من البلوكات كعينة ثم تجمع المعلومات عن الاسر التي تسكن في هذه البلوكات . وعندما نريد سحب عينة من الحيازات الزراعية يمكن أن يكون اطار المعاينة هو قائمة باسماء الحائزين وعناوينهم ويمكن أن يكو قائمة باسماء القرى التي نسحب بعضها كعينة ثم تجمع المعلومات من الحيازات في هذه القرى المسحوبة أو من عينة منها .

وبذلك نلاحظ ان اطار المعاينة يمكن أن يكون قائمة باسماء وحدات أو خريطة مقسمة الى مساحات معينة أو بلوكات معينة ، كما يمكن أن يكون صورة شمسية مأخوذة من الجو (في حالة العينات من الأراضي الزراعية بشرط أن تؤخذ الصور على ارتفاع ثابت دائماً ولذلك لا يمكن استعمال هذه الصور كاطار للمعاينة الا في الدول التي تكون فيها الاراضي الزراعية مستوية غير جبلية) . ومهما كان نوع الاطار يجب أن يشتمل على جميع وحدات المجتمع موضوع البحث وعلى بعض الصفات الاساسية لهذه الوحدات كعناوينها حتى يمكن التعرف عليها عند سحبها في العينة، وكأي صفة اخرى تساعد في التعرف على درجة الاختلاف بين وحدات المجتمع حتى يمكن تحديد نوع العينة التي تناسب هذا المجتمع وعدد الوحدات التي يجب سحبها في العينة .

اية فرصة لان تمثل في العينة ، الامر الذي لا بد ان يؤدي إلى تحيز النتائج الى أعلى ، وتكون الدراسة بذلك دون أية قيمة علمية . ونلاحظ ان عدم شمول الاطار لا يؤدي فقط الى التحيز في سحب العينة وانما يؤدي كذلك الى الخطأ في تقدير النتائج النهائية . فاذا افترضنا أن مجتمع ما يتكون من ١٠٠٠ اسرة مثلاً في حين ان عدد الاسر في قائمة الاطار هي ٩٠٠ فقط (دون ان يعرف الباحث ذلك) وسحب الباحث ١٠٠ اسرة كعينة ، فانه بذلك سوف يقوم بتقدير صفات المجتمع موضوع البحث من العينة على أساس أن كسرها = $\frac{١}{٩}$ بينما يكون الكسر الحقيقي للعينة = $\frac{١}{١٠}$ وبذلك تكون التقديرات متحيزة إلى أسفل . كذلك يمكن أن يكون الاطار غير شامل بسبب قدمه مثل القوائم التي تؤخذ من الادارات الحكومية المختلفة عن المؤسسات الصناعية او التجارية أو غيرها من المؤسسات حيث تكون هذه القوائم في الغالب قديمة فلا تشتمل على المؤسسات الجديدة التي انشأت بعد تاريخ اعداد هذه القوائم . كذلك يمكن ان يكون الاطار غير شامل عند استعمال دليل التلفون حيث لا يكون لدى جميع اسر المجتمع تليفونات ، الأمر الذي يجعل الدراسة متحيزة ضد هذه الاسر وفي صالح الاسر التي يكون لديها تليفونات .

ب - استخدام اطار به ازدواج أي تتكرر فيه أسماء بعض الوحدات وطبعاً يكون التحيز خطيراً اذا كانت هذه الوحدات التي تكررت اسماؤها تمثل نوعاً معيناً من الوحدات له علاقة بموضوع البحث . ويحدث هذا غالبا في القوائم التي تظهر اسماء المحلات التجارية حيث يكون لهذه المحلات فروع مختلفة في احياء مختلفة باسماء مختلفة ، وحيث ان العدادين الذين يقومون بحصر المحلات يوزعون على الأحياء المختلفة لاجراء عملية الحصر فاذا لم يكن لديهم التعليمات بسؤال أصحاب المحلات التجارية عما إذا كان لديهم فروع في أحياء مختلفة ، فان القوائم التي يعدونها سوف

يظهر فيها الازدراج . والتحيز الذي ينتج عن الازدواج يكون بسبب أن الوحدات التي تتكرر اسماؤها تعطى أكثر من فرصة للدخول في العينة بينما لا يكون للوحدات الأخرى إلا فرصةواحدة فقط. كذلك يترتبعلىالازدواج التحيز في التقديرات النهائية ، حيث إذا افترضنـــا ان القائمة باسماء المحلات التجارية تحتوي على ١٠٠٠ وحدة منها ١٠٠ وحدة عبارة عن وحدات متكررة ، أي ان العدد الحقيقي للمحلات التجارية هو ٩٠٠ فقط (دون أن يعرف القائمون على البحث) وسحبنا من هذا الاطار ١٠٠ وحدة كعينة فاننا

في النهاية سوف نقوم بالتقدير على أساس ان كسر العينة = $\frac{١}{١٠}$ بيـــنما

يكون الكسر الحقيقي = $\frac{١}{٩}$ وبذلك تكون التقديرات متحيزة الى أعلى .

جـ – سيطرة الانسان بجسمه أو بعقله على سحب وحدات العينة ، حيث أن الانسان مهما كان محايداً لا بد أن يتأثر بعوامل معينة عند سحب العينة وقد أجريت التجارب الكثيرة التي اثبتت هذه الحقيقة ، فاختيار الانسان لارقام يمكن أن يؤدي الى التحيز حيث انه قد يكره بعض الأرقام لدلالتها السيئة بالنسبة له وقد يحب بعض الأرقام لدلالتها الحسنة بالنسبة له وبذلك لا يمكن أن يكون الانسان عشوائياً صحيحاً في اختيار الأرقام اذا كان مسيطراً على عملية السحب بعقله الخاص . كذلك يكون سحب بطاقات من كيس به بطاقات متساوية الشكل واللون عملا متحيزاً حيث أنه في الغالب تلتصق البطاقات ببعضها ، الأمر الذي يعطي البطاقات الخارجية فرصة الدخول في العينة ولا يعطي البطاقات المضغوطة بينها أية فرصة . وحتى إذا كان عدد البطاقات صغيراً يكون من الأفضل عدم اتباع هذه الطريقة نظراً لاحتمال التحيز مهما كان هذا الاحتمال صغيراً .

د – عدم اجابة بعض الوحدات التي سحبت في العينة أو رفضهم الإجابة

ثم قام العدادون بالاستعاضة عن هذه الوحدات بوحدات أخرى لم يتبعوا في اختيارها أية قاعدة ، وقد سبق أن أشرنا إلى ذلك .

هـ – اتباع طريقة السحب المنتظم إذا كانت وحدات المجتمع مرتبة ترتيباً معيناً له علاقة بموضوع البحث ، مثلا سحب كل عاشر اسم من قائمة باسماء عمال منظمين في مجموعات تتكون كل مجموعة منها من ٩ عمال ورئيس لهم ، وبذلك تكون العينة الناتجة مكونة فقط من رؤساء العمال ومن الواضح أنها بذلك تكون عينة متحيزة . كذلك اذا أخذنا كل خامس اسم مثلا فسوف تكون العينة الناتجة مكونة فقط من العمال أي لن تمثل رؤساء للعمال مطلقاً .

من هذه المناقشة يتبين لنا اننا اذا اردنا ان تكون العينة التي ندرسها غير متحيزة يجب ان يعتمد سحبها على اطار شامل لجميع وحدات مجتمع المعاينة ، كذلك لايجب ان يكون به اي ازدواج، كما لا يجب ان يسيطر الانسان بحسه أو بعقله على سحب العينة ، وكذلك لا يجب ان نتبع طريقة السحب المنتظم اذا كانت وحدات المجتمع مرتبة ترتيباً معيناً له علاقة بموضوع البحث، فاذا تمت عملية السحب دون تحيز يجب ايضاً اعطاء التعليمات المشددة الى العدادين بعدم استعاضة اي وحدة بوحدة اخرى دون اشعار المكتب القائم بالبحث بذلك .

وبذلك يتبادر الى ذهننا السؤال ما هي الطريقة السليمة في سحب العينات ؟ هناك طرق كثيرة تعتمد اساسا على السحب باليانصيب حيث ان الوحدات التي يتم سحبها تظهر عشوائياً دون أي تدخل من جانب ساحب العينة . ولتوفير الجهود التي تبذل في السحب العشوائي الآلي (دواليب اليانصيب مثلاً) قام بعض الاحصائيين باعداد جداول تحتوي على اعداد عشوائية (تم سحبها آليا) مكونة من رقمين ومرتبة في صفوف واعمدة وهي تكافىء تماماً من حيث العشوائية الاعداد التي تحصل عليها لوقمنا بالاختيار العشوائي من بين مجموعة الاعداد ٠٠ ، ٠١ ، ٠٢ ، ٠٣ ، ٠٠،٠٠٠،٩٩ بواسطة دواليب اليانصيب . ومع أن الاعداد مكتوبة في الجداول الخاصة بها في شكل رقمين الا أنه من السهل تحويل هذه الاعداد الى اعداد مكونة من رقم واحد أو ثلاثة أرقام أو

أربعة أو أي عدد آخر من الارقام تستلزمه عملية سحب العينة العشوائية . ونلاحظ أن أي عدد نكونه من هذه الارقام يكون عدداً عشوائياً لأن كل خانة منه عشوائية ووضعها بجانب الاخرى عشوائياً كذلك. وحتى لا يكون للشخص حرية الاختيار اذا تقرر قراءة هذة الأرقامالعشوائية بطريقة هندسية معينة أي في أي اتجاه معين يجب عدم تغيير اتجاه القراءة كلما تراءى لناذلك. كذلك نلاحظ أنه اذا تكرر العدد عند القراءة نستبعده كا انه اذا قرأنا رقماً لا يدخل ضمن أرقام المجتمع موضوع البحث نستبعده أيضاً . وبذلك تكون خطوات العمل كالآتي :

١ ـ وضع حدات المجتمع في قائمة مرتبة حسب الاحرف الهجائية ، واذا كان اطار المعاينة خريطة أو صورة شمسيةلا يكون هناك داعي لهذهالخطوة.

٢ ـ اعطاء هذه الوحدات أرقاماً مسلسلة .

٣ ـ على اساس عدد الارقام في أكبر عدد مسلسل نصل اليه نقرأ الارقام العشوائية ، فاذا كان أكبر عدد مكوّن من رقمين نقرأ الارقـــام العشوائية رقمين رقمين ... وإذا كان أكبر عدد مكون من ثلاث أرقـــام نقرأ الارقام العشوائية ثلاثة أرقام . ويمكن أن نبدأ في القراءة من أي مكان في الجدول على أن نستمر في القراءة حسب نظام هندسي ثابت .

٤ ـ نستبعد الاعـــداد للتي تصادفنا والتي تكون أكبر من أكبر رقم مسلسل لدينا .

٥ ـنستبعد الاعداد التي تصادفنا والتي سبق أن ظهرت في سحب العينة.

٦ ـ اذا كانت الارقام المسلسلة في إطار العينة تصل الى ٢٠٠ مثلاً فاننا نقرأ الأعداد العشوائية من ثلاث خانات ولكننا بذلـــك سوف نقرأ أرقام كثيرة مكونة من ثلاث خانات ولكنها تزيد على ٢٠٠ فنستبعدها ، إلا ان ذلك سوف يضيع علينا وقتاً طويلاً . لذلك نقسم مجموعة الأعداد العشوائية

الى مجموعات من ١ الى ٢٠٠ ومن ٢٠١ الى ٤٠٠ ومن ٤٠١ الى ٦٠٠ ومن٦٠١ الى ٨٠٠ ومن ٨٠١ الى ٩٩٩ ثم نترك المجموعة الاولى كما هي ثم نحوّل كل من المجموعات الباقية الى المجموعة الاولى وذلك بطرح ٢٠٠ من المجموعة الثانية و ٤٠٠ من المجموعة الثالثة و ٦٠٠ من المجموعة الرابعة و ٨٠٠ من المجموعة الخامسة ، وبذلك نكون قد حولنا جميع الأعداد العشوائيـــة المكونة من ثلاث خانات والتي تزيد من ٢٠٠ إلى أعداد عشوائية من ١ الى ٢٠٠ ، فاذا قرأنا مثلاً ٢٣٥ يكون العدد العشوائي الذي نأخذه هو ٣٥ واذا قرأنا العدد العشوائي ٧٢٨ يكون العدد العشوائي الذي نأخذه هو ١٢٨ . وبهذهالطريقة نكون قد أخـــذنا العينة من أعداد متتالية في الجدول ووفرنا الوقت اللازم للاستمرار في القراءة حتى نحصل على العينة المطلوبة .

٧ – كذلك إذا كان لدينا مساكن في شوارع مختلفة وأردنا أن نسحب منها عينة يكفي أن نعرف عدد المساكن في كل شارع حتى نسحب العينةالتي نريدها . نفترض ان توزيع المساكن هو كالآتي :

الشارع	رقم	١	٧٥	مسكن
»	»	٢	١٢	»
»	»	٣	١٩٦	»
»	»	٤	٣٢	»
»	»	٥	٦٣	»
»	»	٦	١٤	»

نكون متجمعاً صاعداً لعدد المساكن في هذه الشوارع كالآتي :

الشارع	رقم	١	٧٥	مسكن
الشارع الاول + الشارع	رقم	٢	٨٧	مسكن
الشارعين الاول + الشارع	رقم	٣	٢٨٣	مسكن

الثلاث الشوارع الاولى + الشارع رقم ٤	٣١٥	مسكن
الاربع الشوارع الاولى + الشارع رقم٥	٣٧٨	مسكن
الخمس الشوارع الاولى + الشارع رقم ٦	٣٩٢	مسكن

وبذلك نعرف اننا اذا رقمنا المساكن ترقيما متسلسلا فسوف يكون آخر عدد هو ٣٩٢ وهو مكون من ثلاث خانات .نقرأ الاعداد العشوائية من ثلاث خانات ونقسمها الى مجموعات تبدأ المجموعة الاولى من ١ الى ٥٠٠ والثانية من ٥٠١ الى ٩٩٩ وتبقى المجموعة الاولى كما هي ونطرح ٥٠٠ من اعداد المجموعة الثانية . نفترض اننا قرأنا العدد العشوائي ٠٨٦ فان هذا العدد يقع في الشارع الثاني ويكون رقم المسكن في هذا الشارع هو رقم :

٨٦ — ٧٥ = ١١ . نفترض اننا قرأنا العدد العشوائي ٧٦٤ نطرح منه ٥٠٠ يكون العدد الباقي هو ٢٦٤ وهذا العدد يقع في الشارع الثالث ورقم المسكن في هذا الشارع هو ٢٦٤ — ٨٧ = ١٧٧ . وهكذا حتى تكون العينة التي نريدها .

٨—نفترض ان لدينا حقل مستطيل طوله ٣٠٠متر وعرضه ٢٠٠متر وأردنا ان نأخذ منه عينة من ثلاث قطع مساحة كل منها متر مربع . نقسم طول الحقل الى امتار تبدأ من الاعداد صفر حتى ٣٠٠ وكذلك نقسم عرض الحقل الى امتار تبدأ من العدد صفر حتى ٢٠٠ . نطرح من الطول متر واحد فيكون آخر عدد على الطول هو ٢٩٩، وكذلك نطرح من العرض متر واحد فيكون آخر عدد على العرض هو ١٩٩ . نقرأ الاعداد العشوائية من ثلاث ارقام فاذا ظهرت اكثر من ٣٠٠ نطرح هذا العدد من العدد الذي قرأناه حتى العدد ٥٩٩ واذا ظهر العدد من ٦٠٠ الى ٨٩٩ نطرح منه ٦٠٠ ويستبعد اي عدد يزيد على ٨٩٩ . وكذلك بالنسبة للعرض . ان قراءتنا الاولى تعطينا عدداً على الطول وقراءتنا الثانية تعطينا عدداً على العرض . عند تقاطع المحورين المقامين من هذين العددين نأخذ مربعاً مساحته متر مربع في اتجاه

الترقيم على كل من الطول والعرض .

ان العينة المسحوبة بهذه الطريقة تسمى بالعينة العشوائية وهي التي يتم سحبها على اساس تكافوء فرص الاختيار لجميع وحدات مجتمع المعاينة – الا ان مثل هذه العينة لها عيوب يمكن ان تظهر بالنسبة لبعض المجتمعات . فالوحدات المأخوذة يمكن ان تظهر على الطبيعة متناثرة تناثراً كبيراً بحيث يصبح العمل في الميدان صعبا جداً وبحيث يكون الاشراف على العدادين اثناء عملهم امراً متعذراً . ولا شك ان هذا العيب يكون بالغ الاهمية اذا كان المجتمع موضوع البحث مترامي الاطراف بحيث يكون تناثر وحدات العينة في هذا المجتمع امراً له أهميته . الا ان العكس يمكن ان يحدث اذ قد تتركز وحدات العينة في قسم من المجتمع ويكون هذا التركز امراً خطيراً حيث يكون عاملا مؤديا الى التحيز اذا كانت اقسام المجتمع تختلف عن بعضها اختلافاً جذريا ، فاذا أردنا مثلا اختيار عينة عشوائية من الاراضي الزراعية في لبنان ونحن نعلم ان اراضيه تنقسم الى اجزاء (اراضي الساحل واراضي الجبل واراضي البقاع) تختلف عن بعضها اختلافات كثيرة من حيث التربة وانواع المحاصيل التي تزرع في كل جزء منها فان هذه العينة يحتمل ان تتركز في اي جزء من هذه الاجزاء ولا تكون بذلك صورة صادقة من هذا المجتمع بل انها بالعكس تكون متحيزة تحيزاً صارخاً . لذلك لا يمكن ان يطمئن ساحب العينة الى العينة العشوائية اذا كان المجتمع موضوع البحث من هذا النوع . كذلك عندما نريد سحب عينة عشوائية لا بد من اعداد اطار شامل لجميع وحدات المجتمع،وهو امر قد يكون من المتعذر في بعض الاحيان نظراً لان اعداد مثل هذا الاطار يحتاج الى حصر وحدات المجتمع حصر شاملا وذلك يحتاج الى تكاليف باهظة قد لا يكون الجهاز القائم بالعمل غير مستعد لتحملها. ومن ناحية اخرى رأينا ان سحب العينة باستخدام الارقام العشوائية فيه ضياع لوقت الساحب في قراءة الأرقام العشوائية خاصة إذا كثرت وحدات المجتمع وكانت العينة

التي نريد سحبها ذات حجم كبير . لذلك قد يكون من الافضل التفكير في نوع آخر من انواع العينات التي فكر فيها علماء الاحصاء لمواجة مثل هذه العيوب في العينة العشوائية البسيطة .

وباستخدام العينة الطبقية يمكن معالجة الاختلاف الكبير بين وحدات المجتمع . ويكون ذلك بتقسيم المجتمع إلى طبقات متجانسة ، ففي مثالنا عن الاراضي الزراعية في لبنان تقسم الاراضي الى ثلاث أقسام – أراضي الساحل وأراضي الجبل وأراضي البقاع . وفي بعض الحالات يكون انقسام المجتمع إلى طبقات أمر موجود فعلاً في الطبيعة كما هو الحال بالنسبة للأراضي الزراعية في لبنان ، إلا انه في حالات اخرى يتحتم على الباحث نفسه اجراء هـــذا التقسيم . وفي مثل هذه الحالات يجب أن يكون لديه صفة معينة أو مقياس معين يمكن على أساسه اجراء التقسيم حيث ان الهدف من التقسيم هو تجميع كل مجموعة من الوحدات المتجانسة في قسم أي طبقة قائمة بذاتها . فاذا كنا ندرس اجور العمال يكون من الأفضل تقسيمهم حسب المهن التي يقومون بها ، وإذا كنا ندرس الانتاج الصناعي يكون من الأفضل تقسيم المؤسسات الصناعية حسب نوع نشاطها وحجم المؤسسات التي تعمل في كل نشاط . بعد تقسيم المجتمع الى طبقات تؤخذ من كل طبقة عينة امـــا بالسحب العشوائي أي باستخدام الاعداد العشوائية أو بالسحب المنتظم الذي سنشير اليه فيما بعد . ثم تدرس كل عينة على حدة ثم تعمم النتائج على المجتمع كله مع ترجيح نتائج كل طبقة بحسب نسبة العينة المأخوذة منها .

بهذه الطريقة نحقق مزايا كثيرة ، فأولاً نؤكد من بادىء الأمر عدم تحيز العينة حيث انها سوف تقسم على طبقات المجتمع المختلفة فلا خوف من تركزها في طبقة دون اخرى كما هو الحال بالنسبة للعينة العشوائية البسيطة ولا شك أن ذلك يؤدي بطريق غير مباشر الى زيادة دقة النتائج . وثانياً يمكننا أن نحصل على صفات كل جزء من المجتمع على حدة بجانب معرفة صفات المجتمع

ككل غير مجزا . ففي مثالنا عن الاراضي الزراعية نستطيع أن نتعرف على صفات الاراضي الزراعية في كل من الساحل والجبل والبقاع بجانب معرفة صفات الاراضي الزراعية في مجموعها بعد ترجيح النتائج الخاصة بكل منطقة على حدة .

وإذا كان لدينا إطار عام بوحدات المجتمع جميعاً مبين به الصفة التي على أساسها سوف نسحب عينة طبقية . فكل ما نريد ان نعرفه هو عدد الوحدات التي نريد سحبها في كل طبقة وبذلك نعد إطارات فرعية لكل طبقة على حدة ، ثم بالقراءة من الاعداد العشوائية إذا ظهرت وحدة ما نعرف صفتها ، وعلى هذا الاساس نضعها في إطار الطبقة الخاصة بها ونستمر في العمل بهذه الطريقة حتى تكون الوحدات المطلوبة في كل طبقة على حدة .

وكما سبق أن ذكرنا يتحدد حجم العينة تبعاً لدرجة الاختلاف بين وحدات المجتمع ، ولهذا اذا كان المجتمع شديد الاختلاف فان العينة العشوائية البسيطة التي يجب سحبها من هذا المجتمع لا بد وأن تكون كبيرة جداً حتى تعطينا درجة دقة معينة . إلا انه بتقسيم هذا المجتمع الى طبقات بحيث تشتمل كل طبقة على الوحدات المتجانسة فان العينة التي يجب سحبها من كل طبقة سوف تكون صغيرة بحيث يكون مجموع الوحدات في العينات التي تؤخذ من الطبقات المختلفة أقل بكثير من حجم العينة العشوائية البسيطة ، وهذه ميزة هامة للعينة الطبقية حيث يمكن على اساسها انقاص تكاليف الدراسة الى أدنى حد ممكن . على انه يجب أن نلاحظ ان هذه الميزة لا تتحقق الا اذا كان تقسيم المجتمع الى طبقات يؤدي الى نقص الاختلاف في كل طبقة على حدة نقصاً كبيراً ؛ أما إذا كان النقص في الاختلاف سوف يكون بسيطاً فان حجم العينة الطبقية يحتمل ان يزيد عن حجم العينة العشوائية البسيطة .

وباستخدام العينات المنتظمة نستطيع أن نختصر كثيراً من الوقت الذي يضيع في السحب باستخدام الأعداد العشوائية . والصفة الاساسية لهذه العينات

هي تساوي المسافات في الاطار بين الوحدات المتتالية التي تؤخذ في العينة ولاستخدام هذه الطريقة في سحب العينات يجب أن نعرف عدد وحدات المجتمع ، وليكن ١٠٠٠ مثلاً فاذا أردنا أن نسحب ١٠٠ وحدة كعينة،فاننا

نقسم حجم المجتمع على حجم العينة $\frac{١٠٠٠}{١٠٠}$ = ١٠ وبذلك نحصل على ما نسميه

مسافة الانتظام وهي ١٠ في هذا المثال . ويعني ذلك انه لا بـــد أن نأخذ وحدة من كل عشرة وحدات من المجتمع . وبذلك نبدأ بقراءة عدد عشوائي حتى العدد عشرة فاذا فرضنا ان هذا العدد هو ٧ يكون العدد التالي لوحدة العينة ١٧ والعدد الثالث ٢٧ وهكذا .

وليست ميزة سهولة السحب هي الميزة الوحيدة للعينة المنتظمة، ففي كثير من الأحيان تكون أصدق تمثيلاً للمجتمع أي أقل تأثراً بخطأ الصدفة من العينات العشوائية البسيطة فأخذ عينة منتظمة من المنازل في أحدى المدن يضمن لنا تمثيل العينة لمنازل الأحياء المختلفة وبنسبة تساوي تقريباً نسبة عدد المنازو في كل حي . وبهذا تكون العينة المنتظمة ادق تمثيلاً في هذه الحالة من العينة العشوائية فيما يختص بالنواحي الاجتماعية والاقتصادية والعمرانية المرتبطة بتوزيع المنازل على الاحياء المختلفة في المدينة .

على انه يجب أن نلاحظ عند سحب العينات المنتظمة عدم انتظام الوحدات في اطار المعاينة انتظاماً له علاقة بموضوع البحث ، وقد سبق أن أشرنا الى ذلك فيما يختص بترتيب العمال في قائمة حيث يمكن أن يكونوا منظمين في مجموعات كل مجموعة منها لها رئيس الذي يظهر اسمه بشكل منتظم في أول كل مجموعة . كذلك لا يجب أن نظن ان العينة المنتظمة ليست الا عينة عشوائية حيث نسحب الوحدة الاولى عشوائياً أى باستخدام الاعداد العشوائية . ولكن يجب الا ننسى أن ليس هناك في العينة تعشوائية ما يمنع من ظهور أي وحدة مهما كان ترتيبها في الاطار ، بينما في العينة المنتظمة اذا تحددت الوحدة

الاولى فان باقي الوحدات فيالعينة تتحددعلىأساس مسافة الانتظام التي نتبعها. كذلك نلاحظ أن عدد العينات التي يمكن سحبها باستخدام السحب العشوائي يكون عدد كبير جداً يصل الى ما لا نهاية ، أما بالسحب المنتظم فان عدد العينات التي يمكن سحبها عدد محدد ، فاذا كانت مسافة الانتظام ١٠ كان عدد العينات التي يمكن سحبها عشرة عينات فقط واذا كانت مسافة الانتظام ٢٠ فان عدد العينات التي يمكن سحبها عشرين عينة فقط . ولذلك نقابل في العينات المنتظمة صعوبات تحليلية حيث أن دراسة خطأ الصدفه على النتائج أصعب عادة في هذه الحالة منه في حالة العينات العشوائية .

وباستخدام العينات على مراحل نستطيع ان نتغلب على صعوبة اعداد اطار للمجتمع كله خاصة اذا كان هذا المجتمع واسع جداً بسبب الصعوبات المالية والادارية التي نواجهها في هذه الحالة . ففي بحث عن الحيازات الزراعية قد نجد انه من العبث أن نفكر في أخذ عينة عشوائية أو متنظمةفي الدولة بأجمعها . وفي مثل هذه الحالة يتعين علينا تجميع العينة في مناطق محدودة حتى نستطيع أن نتغلب على الصعوبات المالية والادارية، ويكونذلك بأخذعينة على مراحل وفي كل مرحلة نصغر وحدةالعينة عنالمرحلة السابقة لها حتى نتنهي بوحدة البحث في المرحلة الأخيرة . ففي المرحلة الأولى نأخذ عينة من الأقضية ومن عينة الاقضية نأخذ عينة من القرى ومن القرى نأخذ عينةمن الحيازات الزراعية التي نريد دراستها .

٦ – تقدير الميزانية الزمنية والمالية والبشرية الضرورية للبحث –

إن عمليات بحوث الميدان تكون من التداخـــل والتشابك بحيث لا بد من إعداد جدولاً زمنياً تبين فيه العمليات المختلفة التي يتكون منها العمل وتوضع أمام كل عملية خط يبين متى يجب أن تبدأ ومتى يجب أن تنتهي وبذلك يبين المدة اللازمة لانهائها . والهدف الأساسي من هذا الجدول أن يطمئن الباحث في أي وقت يشاء الى ما إذا كان يستمكن من الانتهاء من بحثه ونشر نتائجه

في الموعد المحدد ، ولـــكي يعرف مواطن التأخير فيعمل على تعجيلها حتى لا تتراكم التأخيرات وتؤدي في النهاية إلى عجز الباحث عن المحافظة على المواعيد المقررة . ولا بد عند اعداد هذا الجدول أن نراعي كل العوامل التي يمكن ان تؤثر في سير العمل كالمرض والفصل والاجهاد والعطلات وظروف الطقس وما إلى ذلك من عوامل يمكن ان يكون لها تأثير مهما بلغ الباحث من الدقة في التصميم والاحكام .

كذلك يجب أن يلتزم الباحث ببرنامج انفاقي حتى يطمئن الى امكان اتمام البحث والسير فية حتى نهايته . ولا يعني ذلك تخصيص مبلغ معين للبحث وانما المقصود هو اعداد ميزانية مفصلة تظهر المبالغ المقدرة لكل عملية من العمليات التي يتكون منها البحث . وفي الغالب تعد ميزانيات البحوث على أساس تقسيم العمل الى ثلاث مراحل - مرحلة الاعداد للنزول الى الميدان وتتضمن وضع خطة البحث وتصميم الاستمارات اللازمة وصياغة التعليمات والارشادات وتجنيد وتدريب العدادين والمشرفين على الأعمال التي سوف توكل اليهم والقيام بالبحث التجريبي ، ومرحلة النزول الى الميدان لجمع المعلومات ومراجعتها ، ومرحلة التبويب وتشمل عمليات الترميز والتثقيب والفرز والتبويب وتحليل البيانات و كتابة التقرير ونشره ان أعداد ميزانية مفصلة لهذه العمليات المختلفة يساعد في التوصل الى تقدير للرقم الاجمالي للتكاليف دون أن يكون ذلك مبنياً على الحدس والتخمين . كذلك تساعد الميزانية التفصيلية في مراقبة الانفاق على العمليات المختلفة للتأكد من عدم التلاعب في ذلك . واعداد الميزانية التفصيلية لا يعني تقدير مبالغ جامدة لا يمكن تغييرها فيما بعد ، حيث يمكن أن تستجد ظروف مثل ارتفاع أسعار الورق تؤدي الى ضرورة التغيير . لذلك بعد أن ينتهي العمل لا بد من أعداد حساب ختامي يظهر ما أنفق فعلاً على كل عملية من العمليات حتى يكون هذا الحساب أساساً سليما نرتكز عليه عند اعداد ميزانيات البحوث في المستقبل .

وبناء على القرارات السابقة فيما يختص بالمعلومات التي يرغب الباحث في

جمعها ، ونطاق البحث وشموله ، ومجال جمع المعلومات ، والجدول الزمني يستطيع الباحث أن يقدر الميزانية البشرية أي عدد الموظفين الضروريين للقيام بكل عملية يتطلبها البحث .

ثانياً ــ تصميم الاستمارة والتعليمات الخاصة بها :

عند تصميم الاستمارة الخاصة بموضوع البحث يجب ملاحظة الآتي :

١ــ لا يجب أن تشتمل الاستمارة الا على الأسئلة التي لها علاقة مباشرة بموضوع الدراسة ، ولذلك يجب الموازنة بين الفائدة التي تعود من اضافة سؤال ما والزيادة في التكاليف والوقت والجهود التي تنشأ عن ذلك ، بحيث اذا تبين ان الفائدة منه لا تتناسب وزيادة التكاليف وجب اهماله وعدم ادراجه في الاستمارة . ولكي يتأكد الباحث أن الاستمارة ستكون شاملة لجميع البيانات اللازمة لتحقيق أغراض الدراسة يجب اعداد الجداول التخيلية التي سوف تصنف وتبوب فيها المعلومات بعد جمعها . وعند التفكير في أسئلة الاستمارة يجب أن تكون هذه الجداول في مخيلة المصمم دائماً بحيث يستطيع أن يحدد كيف ستظهر اجابة كل سؤال كبيان مصنف في واحد من هذه الجداول أو في بعضها . وبهذه الطريقة في التصميم يمكن أن يتقرر اهمال بعض الأسئلة أو اعادة صياغة بعضها أو التفكير في أسئلة اخرى .

٢ ــ لايجب أن تحتوي الاستمارة ألا على الأسئلة التي يكون جميع المستجوبين أو غالبيتهم على الأقل قادرين على الاجابة عليها ، حيث يجب أن يؤخذ عامل الزمن في الاعتبار اذا يمكن أن تكون الاجابة قد توفرت لدى المستجوب في وقت ماض الا أنه نسيها بفعل مرور الزمن فلا يصح مثلاً أن نسأل عن دخل المستجوب قبل أعوام مضت أو عدد ساعات العمل التي عملها في الماضي .

٣ ــ يجب تجنب الأسئلة التي يكون فيها حرج للمستجوب نظر لعلاقتها باحواله الشخصية ، وبذلك يجب الابتعاد عن الأسئلة التي يعتبرها المستجوب

تدخلا في حياته الخاصة أو لها مساس بتقاليده وقيمة الأخلاقية . واذا اضطر الباحث الى مثل هذا النوع من الأسئلة الشخصية يحسن توظيف عدادين ليسوا من أهل المنطقة التي يعملون فيها، كما يحسن ادراج هذه الأسئلة بين أسئلة اخرى كثيرة ليس لها الصفة الشخصية .

٤ – يجب تجنب الأسئلة التي تستدعي الاجابة عليها اجراء عمليات حسابية من قبل المستجوبين اذ غالباً لا يعطون اجابات دقيقة عليها . فالسؤال عما تنفقه الاسرة شهرياً على الملابس والمصروفات المدرسية يستدعي اجراء عمليات حسابية نظراً لان هذه النفقات تكون في الغالب فصلية وليست شهرية .

٥ – يجب تجنب الأسئلة الايحائية أي تلك التي توحي باجابة معينة يريد الباحث اثبات صحتها ، ومن الأسئلة على ذلك – هل تغيبت بسبب المرض ؟ هل تأخرت بسبب سوء المواصلات ؟ هل سرقت لشعورك بالجوع ؟ والاجابة على مثل هذه الأسئلة يكون في الغالب بالايجاب .

٦ – يجب اختيار الالفاظ البسيطة السهلة ، كما يحسن اختيار الشائع منها بين المستجوبين، لذلك يجب أن يؤخذ المستوى التعليمي للمستجوبين في الاعتبار عند وضع الأسئلة . ان الفاظا مثل الرسوم الجمركية ، الضرائب غير المباشرة ، اليساريين ، الحالة الزواجية ، النشاط الاقتصادي ، البطالة ، ... الخ قد تبدو واضحة في ذهن الباحث ولكنها بالتأكيد تكون غير مفهومة بتاتاً من بعض المستجوبين وغير واضحة المعنى بالنسبة للبعض الآخر . واذا اضطر مصمم الاستمارة الى استخدام هذا النوع من الألفاظ يجب أن يشرح معناها في التعليمات المفسرة للاستمارة .

٧ – لا يجب أن تكون الأسئلة طويلة، إذ قد تكون الألفاظ التي يتكون منها السؤال بسيطة للغاية الا أن المستجوب لا يدرك المقصود منها نظراً لطول السؤال وشموله لمقاطع كثيرة . ان المستجوبين غالباً يطالبون العدادين باعادة

قراءة مثل هذه الأسئلة مراراً ، الأمر الذي يؤدي الى اضاعة الوقت ، وفي بعض الأحيان يكون طول السؤال يسبب ادماج سؤالين أو أكثر في سؤال واحد ، الأمر الذي يؤدي الى ارتباك المستجوب وارتباك العمل في المكتب عند اجراء التبويب ، فسؤال مثل هل انت عضو في أحد النوادي الرياضية وما هي الرياضة التي تمارسها وأيها المفضلة عندك ؟ هذا السؤال يجب تقسيمه الى ثلاث أسئلة فيصبح كل منها قصيراً وبسيطاً وواضحاً ، الأمر الذي يساعد العداد على القاء كل سؤال منها والمستجوب على فهمها بسرعة والاجابة عليها بدقة والمكتب على اجراء التبويب بسهولة ويسر .

٨ – يجب تجنب الأسئلة النسبية ، مثلا هل تذهب الى السينما كثيراً ؟ سؤالى يترتب عليه اجابات بمعايير مختلفة حيث أن ما يعتبره البعض كثيراً يعتبره البعض الآخر قليلا .

٩ – يجب تجنب الألفاظ المثيرة مثل الرأسماليّة والشيوعية والمحافظين والأحرار واليساريين ورجال الأعمال والطبقة العاملة .. الخ ، حيث يكون لمثل هذه الألفاظ وقع مختلف على المستجوبين ، الامر الذي يوجد تحيزاً في اجاباتهم . كذلك لا يجب استخدام الاسماء الضخمة مثل هـــل توافق على رأي ... ؟ حيث أن عاطفة المستجوبين نحو هذه الاسماء تختلف ، الامر الذي يوجد تحيزا في اجاباتهم . كذلك يجب الابتعاد عن الألفاظ الايحائية مثل كلمة « معقولة » حيث انه اذا استخدمت هذه الكلمة في سؤال مثل هل توافق على زيادة معقولة في الأسعار بأمل أن يؤدي ذلك الى الانتعاش ؟ على اساس هذا السؤال سوف يجيب نسبة كبيرة من المستجوبين بنعم وتقـل هذه النسبة كثيراً اذا حذفت كلمة معقولة من السؤال .

١٠ – يجب الحذر من استخدام الجمل التي يمكن أن يكون لها ارتباط بضمير المستجوب أو شعوره بالكرامة والفضيلة حيث ان غالبية الناس يرغبون في الشعور بانهم أذكياء، منطقيين ، كرام، صالحين، محبين للغير ، وهم لذلك يميلون الى الاجابة على اساس ما يجب أن يكون سلوكهم أو رأيهم بالنسبة لوضع معين وليس على أساس سلوكهم أو رأيهم الواقعي .

١١ – كذلك يجب تجنب الأسئلة التي تدفع المستجوب الى الادعاء ، مثلا السؤال هل قرأت كتاب كذا ؟ يمكن الاستعاضة عنه بالسؤال هل تنوي قراءة كتاب كذا ؟

١٢ – يحسن أن تتضمن الاستمارة بضع أسئلة يكون الغرض منها تدقيق اجـــابات الأسئلة الأخرى حيث يمكن السؤال عن نفس الشيء بطريقتين مختلفتين ووضع السؤالين في مكانين مختلفين من الاستمارة حتى يساعدنا ذلك في تدقيق التماسك الداخلي للإجابات ... فمثلا يمكن التأكد من صحة الاجابة على سؤال عن العمر اذا وضع سؤال آخر عن العمر عند الزواج ومـــدة الحياة الزواجية او عن تاريخ الاقتراع للجندية . وفي استمارة ميزانية الأسرة قد لا يكون من أغراض الدراسة جمع معلومات عن الدخل والاقتراض والادخار حيث يكون الغرض معرفة نفقات الاسرة فقط ومع ذلك يحسن إضافة أسئلة عن هذه العناصر الثلاث حتى تكتمل الصورة التي نأخذها عن الاسرة فيمكن بـــذلك تدقيق اجاباتها عن الانفاق .

١٣ – يحسن تحليل السؤال الى عناصره المختلفة ، مثلا إذا كنا نسأل عند تفضيل المستجوب لنوع معين من السيارات يجب أن نتذكر ان هناك عوامل كثيرة للمفاضلة كالتكلفة والاداء والحجم والمظهر ومدى توفرها ، وكل من هذه العوامل له أيضاً جزئيات ، فالتكلفة مثلا تنقسم الى قسمين قسم يتمثل في ثمن شراء السيارة وقسم آخر يتمثل في النفقات الجـــارية لاستخدامها . وبذلك يكون إجمال الاسئلة عن المفاضلة في سؤال واحد مؤدياً في الغالب الى إجابات مضللة .

١٤ – يجب صياغة الأسئلة بحيث لا تستدعي الاجابة عليها الا أقل ما يمكن من الكتابة ، وبمعنى آخر ان تكون الاجابة عليها قابـــلة للتبويب في الجداول المعدة بكل سهولة ويسر ، ولذلك يحسن صياغـــة الاسئلة بحيث تكون الاجابه عليها باحدى الطرق الآتية :

١ – كتابة نعم أو لا . – ٢ – كتابة رقم معين . – ٣ – كتابة علامة معينة في مربع صغير خاص باجابة معينة . – ٤ – وضع خط تحت الاجابة الصحيحة . – ٥ – إحاطه الاجابة الصحيحة بدائرة . – ٦ – كتابة رمز معين . – ٧ – كتابة كلمة أو جملة بسيطة .. على ان صياغة الاسئلة بحيث تكون الاجابة عليها باحدى الطرق الواردة في الفقرات من ٣ الى ٥ تحتاج الى حصر جميع الاجابات المحتملة على كل سؤال بحيث لا يمكن أن تأتي إجابة أي من المستجوبين مختلفة عن الإجابات التي تقرر طبعها في الاستمارة . لهذا يجب التفكير في كل الاجابات المحتملة لكل سؤال دون أن ننسى احتمال الاجابة بـ « لا أعرف » أو « أشك » أو « لا أذكر » أو « لا رأي لي » أو « لا أستطيع الاختيار » . أما إذا رفض المستجوب الاجابة فيسجل العداد ذلك على السؤال دون الحاجة الى طبع هذه الاجابة المحتملة حيث ان طبعها في الاستمارة يشجع العداد على عدم بذل أي مجهود للحصول على إجابة معينة .

هذه بعض القواعد والشروط التي يجب ملاحظتها عند تصميم استمارة البحث ، واحب أن اشير إلى أن ادراك هذه القواعد لا يعني اننا تستطيع أن نصمم استمارة ما تصميما جيداً حيث أن هذا العمل يحتاج إلى خبرة طويلة ومعرفة صحيحة بعلم النفس وعلم الاجتماع . إلا أن ذلك لا يجب أن يخيفنا كثيراً حيث أن كثيراً من الاستمارات الاحصائية لا نواجه في تصميمها صعوبات كثيرة وهي تلك التي ترتبط بالبحوث الموضوعية مثل البحوث الخاصة بالصناعة أو الزراعة أو التجارة ... الخ حيث لا يكون لرأي المستجوب أي اعتبار في الاجابات التي نريدها ، إذ تتطلب الأسئلة في هذه البحوث الإجابة فيشكل أرقام مقيدة في سجلات المستجوبين ، وحتى إذا لم تكن مقيدة لا يكون هناك أية صعوبةفي الحصول عليها إذا أمكن اقناع المستجوبين بسرية المعلومات التي يعطونها وبأنها لن تستخدم ضدهم بأي حال من الأحوال وفقاً لقانون الاحصاء الذي سبق الاشارة اليه .

بعد صياغة أسئلة الاستمارة تكون الخطوة التالية هي تجميعها في مجموعات من حيث الترابط فيما بينها ، فاذا كنا بصدد بحث عن الصناعة مثلا ، يجب أن تكون الأسئلة الخاصة بعدد المشتغلين في المؤسسة منفصلة عن الاسئلة الخاصة بالاجـــور النقدية والعينية التي دفعت لهم وهذه منفصلة عن الأسئلة الخاصة بالانتاج وهكذا . وفي داخل كل مجموعة يجب ان نلاحظ التسلسل المنطقي للأسئلة حتى لا تنقطع سلسلة أفكار المستجوب اثناء اجابته . فلا يجب مثلاً السؤال عن المستوى التعليمي للشخص إلا بعد السؤال عما إذا كان متعلماً او امياً ، كما لا يجب السؤال عن مهنته إلا بعد السؤال عما إذا كان مشتغلا أو متعطلا . كذلك يجب ترك فراغ مناسب لكتابة الاجابة عن كل سؤال على حدة وفراغ آخر لكتابة الرموز اذا كان التبويب للاجابات سوف يجري آلياً .

ولا يكتمل تصميم الاستمارة إلا باعداد التعليمات التي تفسر معاني الكلمات المختلفة التي استخدمت في الاسئلة والتي يمكن ان يساء فهمها من قبل المستجوبين ، ومعنى كل سؤال على حدة حتى يمكن أن يفهمه المستجوب نفس الفهم الذي يفهمه الباحث ، والتي تشرح كيفية الاجابة على كل سؤال ومتى تكون الاجابة صحيحة ومتى تكون خاطئة، ثم كيفية كتابة الاجابة في الفراغ المخصص لها . فاذا كانت الاستمارة التي سوف نستخدمها هي صحيفة استبيان ترسل بالبريد للمستجوبين يحسن طبع التعليمات الخاصة بكل سؤال اسفله مباشرة ولكن باحرف أصغر وفي جمل مركزة قصيرة وذلك بجانب أعداد خطاب يرفق مع الاستمارة لتوضيح الغرض من البحث وسرية البيانات وعدم استخدامها ضد صالح المستجوب ورجاء ملأ الاستمارة واعادتها إلى الباحث في تاريخ أقصاه كذا . أما إذا كانت الاستمارة التي سوف نستخدمها كشف البحث ، ففي هذه الحالة تطبع جميع التعليمات سواء الخاصة بالبحث في مجموعة من ناحية الغرض منه والسرية وسواء الخاصة بتفسير وتوضيح الاستمارة أو الخاصة بمقابلة المستجوب وكذلك الخاصة

بالنواحي الادارية لعمل العداد في الميدان في كتيب مستقل يوزع على العدادين ليسترشدوا به عند قيامهم بعملهم

ثالثاً : تجنيد وتدريب العدادين :

إذا تقرر استخدام كشف البحث يكون الجهاز القائم بالعمل في حاجة إلى عدادين يمكن أن يزيد عددهم إلى الآلاف في الدراسات بالعد الشامل ، الأمر الذي يدعو إلى تجنيدهم من الطلبة والمدرسين وقت العطلات المدرسية ، ويمكن أن يقل عددهم كثيراً في الدراسات بالمعاينة وبذلك يمكن الاعتماد على موظفي الجهاز الاحصائي الذين لديهم خبرة سابقة بهذا العمل .

ومهما كان عدد العدادين الذين يكون البحث في حاجة اليهم ، فان نجاح البحث يتوقف أولاً وقبل كل شيء على استخدام العدادين الصالحين للعمل الذي سوف يوكل اليهم . ولهذا يجب ملاحظة توفر الشروط الآتية في توظيف العدادين :

١ – القدرة على التحدث بطلاقة مع طوائف المجتمع المختلفة .

٢ -- القدرة على جذب انتباه الناس بسرعة وبطريقة سليمة .

٣ – الملاحظة السريعة وادراك التفاصيل .

٤ – الاصرار على النجاح في العمل واتمامه بشكل كامل دون أي نقص

٥ – حب الناس ومعاملتهم بالعطف والطيبة .

٦ – الضمير الحي والأمانة وتحمل المسؤولية .

٧ – الافق الواسع دون الظهور بمظهر الشديد الذكاء .

٨ – الذاكرة الجيدة .

٩ – حب البحث – العقلية الباحثة .

١٠ – القدرة على استيعاب تعليمات كثيرة واتباعها بدقة .

١١ – القدرة على تلخيص ما يسمع من اجابات وتسجيلها تسجيلا موضوعياً .

١٢ – تجنب التحيز الشخصي .

١٣ – الصحة الجيدة والطاقة الجسدية الفائقة .

١٤ – المظهر والسلوك الحسن .

١٥ – الخط الواضح .

١٦ – أن يكون حراً من أية مسئوليات أخرى بحيث يمكنه السفر إذا استدعى الأمر وبحيث يكون راغباً وقادراً على العمل في المساء وفي أيام الاجازات .

١٧ – أن يكون حائزاً على مؤهل علمي يتفق ومستوى الدراسة .

١٨ – يتوقف جنس العداد على نوع الدراسة ، ففي الدراسات التي تتعلق بربات المنازل يحسن أن يكون العدادين من الاناث بينما يحسن توظيف عدادين من الذكور في الدراسات التي تتعلق بغير ربات المنازل .

١٩ – من الأفضل أن يكون عمر العداد بين ٢٥ و ٤٠ سنة حيث أن صغار السن ينقصهم القدرة على الاتصال بالناس بنجاح بينما لا يميل كبار السن إلى اتباع التعليمات بدقة بالاضافة إلى عدم ميلهم إلى الالحاح اللازم لحفز المستجوبين على قبول مقابله العداد والاستماع اليه واعطائه الاجابات المطلوبة.

٢٠ – في الدراسات الواسعة النطاق والتي تشمل أنحاء الدولة المختلفة تنشأ مشكلة تجنيد العدادين محلياً (من كل منطقة للعمل فيها) أو مركزياً (توزيع العدادين على المناطق المختلفة دون التقيد بمنطقة العداد نفسه) .

ولكل من الطريقين مزاياها وعيوبها ، فالتجنيد المحلي يساعد على ضمان معرفة العداد للجهة التي يعمل فيها ولعادات أهلها وتقاليدهم، إلا أن المستجوبين يميلون غالباً إلى عدم اعطاء معلومات عن أنفسهم لاشخاص من منطقتهم خوفاً من استخدام هذه المعلومات مادة للحديث في المقاهي والمجالس العامة ، ويظهر هذا الاتجاه بشكل واضح في الدراسات التي يسأل فيها عن معلومات يعتبرها المستجوبون أسراراً بالنسبة لهم . أمـــا التجنيد المركزي فلا يحقق مزايا التجنيد المحلي التي سبق الاشارة اليها إلى ان المستجوبين لن يتخوفوا من افشاء العدادين لأسرارهم ، الأمر الذي يشجعهم على التعاون معهم . والمفاضلة بين الطريقتين تتوقف إلى حد كبير على نوع الدراسة فإذا لم تشتمل الدراسة على أسئلة عن معلومـــات سرية بالنسبة للمستجوبين يكون من الأفضل اتباع الطريقة الأولى مثل تعداد السكان ، أما إذا اشتملت على مثل هذه المعلومات يحسن اتباع الطريقة الثانية مثل التعداد الزراعي .

على أن التقيد بالشروط السابقة يتوقف الى حد كبير على عدد العدادين اللازم للعمل في الميدان ، حيث انه اذا كان العدد كبيراً فان الهيئة المشرفة على الدراسة سوف تجد نفسها مصطرة الى التساهل في تطبيق هذه الشروط حتى تضمن حصولها على العدد المطلوب خاصة في الدول التي تجري فيها مثل هذه الدراسات لأول مرة . أما في الدراسات بالمعاينة فان الهيئة المشرفة على البحث يجب أن تتقيد الى حد كبير بالشروط السابقة لتضمن نجاح الدراسة حيث أن تساهلها سوف يزيد من نسبة الأخطاء العامة التي يمكن أن تظهر بالاضافة الى أخطاء المعاينة نفسها .

ويتوقف عدد العدادين الذين يحتاجهم البحث على العوامل الآتية . —

١ — عدد وحدات الدراسة ٢ - التوزيع الجغرافي لهذه الوحدات . ٣ — حالة المواصلات في الدولة . ٤ - طول الاستمارة أو قصرها . ٥ — نوع المعلومات المطلوب استعمالها (كشف بحث أو صحيفة استبيان) ، حيث في

حالة استخدام صحيفة استبيان يكون الجهاز في حاجة كذلك الى عدادين للاحقة المستجوبين الذين لا يعيدون استماراتهم الى المكتب في الموعد المحدد لذلك .

واذا ما تم تجنيد العدادين الذين يحتاجهم البحث يبدأ تدريبهم ، ويتوقف نجاح برنامج التدريب على الذين سوف يوكل اليهم تدريب العدادين والاشراف عليهم في الميدان . لهذا يجب عند توظيف المشرفين التقيد بالشروط السابقة فيما يتعلق بالعدادين دون أي تساهل بالاضافة الى ضرورة أن يكون المشرفين من مستوى علمي أعلى ولديهم قدرة أكبر على الاستيعاب والبحث والملاحظة ومواجهة المشاكل على أن الصفة الأساسية التي يجب أخذها بعين الاعتبار عند توظيف المشرفين هي القدرة على قيادة الناس وتوجيههم وكسب احترامهم دون عنف بجانب القدرة على توضيح الامور وشرحها .

ويمكن أن يتخذ برنامج التدريب الخطوات التالية :

١ – اعداد المشرفين والقائمين على تدريب العدادين .

٢ – توزيع العدادين في حلقات للدرس (محلياً أو مركزيا) .

٣ – اعداد جميع المطبوعات اللازمة للتدريب وطبعها حتى يمكن توزيعها على العدادين وقت بدء التدريب . ويجب أن يتوفر لدى كل عداد نسخة من المحاضرات التي سوف تلقى اثناء التدريب ونسخة من كتيب التعليمات. أما غير ذلك من المطبوعات الادارية مثل كشوفات حصر وحدات الاستبيان في مناطق العد المختلفة ونماذج تقارير العمل اليومي فتوزع على العدادين عندما يتقرر قيامهم بالعمل . وقبل هذا أو ذاك يجب أن يكون لدى العداد بطاقة هوية تثبت وظيفته كعداد .

٤ – القاء محاضرات تشرح التعليمات المختلفة التي تتعلق بعمل العدادين في الميدان . وتشمل هـــذه التعليقات كيف يمكن ان يكون العداد ناجحاً في

مقابلة المستجوب ، وشرح للاسئلة المختلفة التي تتضمنها الاستمارة ، الغرض من كل سؤال ومعناه وكيف يمكن توجيهه الى المستجوب وكيفية الحصول على الاجابات الصحيحة وكيفية تسجيل هذه الاجابات في الاماكن المخصصة لها، وكيفية مواجهة المشاكل المختلفة التي يمكن ان تعترض العداد اثناء عمله،وكيفية اجراء مراجعة سريعة لما كتبه في الاستمارات قبل مغادرة مكان المستجوب، وكيفية كتابة التقارير اليومية عن عمله،والاجراءات الادارية التي يتطلبها عمله في الميدان مثل استلام الاستمارات وعدها وكيفية الاتصال بالمشرفين عليهم وكيفية تسليم الاستمارات بعد الانتهاء منها،وكيفيةمواجهة النفقات التي يحتمل ان يحتاج الى انفاقها اثناء متابعة العمل .

ومناقشة هذه التعليمات تكون بشرحها واعطاء أمثلة كثيرة تكشف عن الصعوبات المختلفة التي يحتمل ان تظهر عند العمل بحيث تؤدي المناقشة الى فهم العدادين للتعليمات فهما دقيقاً وكاملا واستيعابها بحيث تصبح شيئاً عادياً بالنسبة لهم وبحيث يكونوا مستعدين للاختبار فيها في أي وقت دون سابق انذار .

٥ – اختبار العدادين في نهاية التدريب ويتوقف استمرار العداد في عمله على نتيجة هذا الاختبار .

٦ – تدرب العدادين الذين اجتازوا الأختبار النظري في الميدان والقصد من هذا التدريب التحقق من أن العداد قد استوعب فعلاً التعليمات التي درسها نظرياً وأصبح بذلك قادراً على تطبيقها عملياً . كذلك يترتب على هـــذا التدريب ان يتعرف العداد على منطقة العد المخصصة له– حدود المنطقة والوحدات الواقعة فيها وأماكن وجود هذة الوحدات (أين تقع ابواب المساكن مثلا)–،ويكون هذا التعرف على أساس مطابقة ما هو مسجل في القوائم الموزعة عليهم مع ما هو موجود على الطبيعة . ويمكن أن يجري هذا النوع من التدريب أثناء اجراء العد الاختباري الذي سيأتي مناقشته فيا بعد.

رابعاً – حصر وحدات المجتمع موضوع البحث :

في دراسة واسعة النطاق كتعداد السكان ، او التعداد الصناعي، او التعداد الزراعي ... الخ . يكون من الضروري تقسيم الدولة الى مناطق عد يخصص كل منها لعداد واحد أو لفريق من العدادين . ويستلزم هذا العمل تحديد أقسام الدولة علي خرائط كبيرة للمناطق الريفية وعلى تصميمات هندسية للمدن . فاذا لم تكن هذه الخرائط والتصميات متوفرة يكون من الضروري العمل على اعدادها خصيصاً لهذه المناسبة .

ويساعد العد الاختباري في تحديد الوقت الكافي لملأ الاستمارات وهذا بالتالي يمكن أن يساعدنا في تحديد مساحة مناطق العد . وهذه المسألة من الأهمية بمكان كبير في تعداد السكان حيث يجب أن تجري عملية التعداد خلال وقت قصير جداً . فاذا كانت مناطق العد واسعة جداً فان عملية التعداد لا بد وان تنهار ، أما اذا كانت ضيقة جداً فان العدادين سوف يجدون أنفسهم بدون عمل خلال جزء من الوقت المحدد وبذلك نكونقد أسرفنا في الانفاق.

بعد تحديد مناطق العد تعد قوائم باسماء وحدات الاستبيان في كل منطقة منها لتوزيعها على العدادين . وفي التعداد الصناعي تتخذ هذه القوائم شكل كشوفات باسماء المؤسسات الصناعية وعناوينها ، وفي التعداد الزراعي تتخذ شكل كشوفات باسماء الحائزين وعناوينهم ، وفي تعداد السكان تتخذ شكل كشوفات باسماء وحدات التعداد التي قد تكون اسرا أو مؤسسات سكنية . ونلاحظ انه في هذا التعداد لا يكفي فقط اعداد هذه الكشوف بل يجب أيضاً ترقيم المساكن في المدن والقرى واعداد علامة مميزة يمكن وضعها عليها عند الانتهاء من العد بالنسبة لها .

ان العناية بعملية حصر وحدات الاستبيان واعداد كشوفات بها تحدد لنا ما إذا كان التعداد سوف يكون شاملاً أو جزئياً . وتظهر أهمية الحصر في تعداد السكان حيث لا يكون من المنطق ترك عملية تحديد موقع وحدات

العد للعدادين كي يقوموا بها اثناء قيامهم بعملية العد حيث يكون الوقت المحدد لهذه العملية ضيق جداً وبذلك لا يمكن أن يضيع العداد وقته بحثاً عن وحدات الاستبيان التي يطلب اليه زيارتها . ولذلك يجب أن يتم هذا العمل مقدماً حتى يستطيع كل عداد أن يتعرف على وحدات العد التي تدخل في منطقة العد الخاصة به قبل الوقت المحدد للتعداد .

أما في أنواع التعداد الأخرى قد يكون من الممكن أن نجمع بين تحديد وحدات العد وزيارتها في نفس الوقت ، وبذلك يمكن تلافي النفقات التي يتحملها الجهاز الاحصائي في اعداد قوائم الحصر ، خاصة إذا تذكرنا أن العد في هذه التعدادات يكون على مراحل . الا انه من الأفضل اجراء الحصر قبل بدء التعداد حتى يمكن تدقيق عمل العدادين على الاقل ، بالاضافة إلى انه يمكن الاسترشاد بقوائم الحصر في تحديد عدد الاستمارات بدقة حتى نتجنب الاسراف . ومن ناحية أخرى قد يكون من المقرر اجراء التعداد لنوع معين من وحدات المجتمع الاحصائي موضوع البحث وبذلك تكون عملية الحصر ذات أهمية كبيرة حيث يمكن خلال هذه العمليـــة معرفة الصفة الخاصة بوحدات المجتمع والتي نسترشد بها في تحديد الوحدات التي يشملها التعداد وتلك التي لا يشملها، مثلاً اذا تقرر أن يشمل التعداد الصناعي المؤسسات التي توظف ١٠ عمال وأكثر ، ففي هذه الحالة لا بد أن نعرف توزيع المؤسسات تبعاً لعدد المشتغلين فيها مقدماً حتى نقتصر فيما بعد أثناء عملية العد على المؤسسات المطلوبة إذ يطلب إلى العدادين زيارة هذه المؤسسات مباشرة تبعاً لعناوينها التي تظهر في كشوفات الحصر .

كذلك في الدراسات بالمعاينة يكون من الواجب حصر وحدات المجتمع لاعداد قوائم باسمائها وعناوينها وبعض المعلومـــات المبدأية البسيطة عنها . وتساعدنا مثل هذه القوائم على تحديد اطار المعاينة (ويجب أن يكون شاملا دون أي ازدواج حتى لا تكون العينة متحيزة) ، وعلى تحديد

نوع العينة الذي يناسب المجتمع موضوع البحث إلا انه قد يتقرر سحب العينة على أساس وحدات كبيرة (العينة على مراحل) أو يتقرر أن يكون الاطار في شكل خريطة أو صور شمسية ، وفي هذه الحالات لا يكون هناك داعياً لحصر جميع وحدات المجتمع وانما نكتفي بحصر الوحدات الكبيرة ثم الوحدات الصغيرة في عينة المرحلة الاولى فقط (حصر الحيازات في عينة القرى مثلاً) أو باعداد الخرائط أو التصاميم او الصور الشمسية التي يكون قد تقرر استعمالها كاطار للمعاينة .

خامساً - تنظيم الدعاية :

مهما حاولنا أن نعد العدة للبحث الاحصائي الذي تجمع معلوماته من الجمهور فان نجاحه يتوقف في النهاية على استجابة المستجوبين بالتعاون في اعطاء المعلومات الصحيحة . ولهذا فالتنظيم المحكم لحملة من الدعاية لكسب ثقة الجمهور وتأييده يكون جزءاً هاماً من العمل التمهيدي .

ويرتكز جزء هام من كسب ثقة الجمهور وتعاونه على العدادين أنفسهم ، ولهذا فتدريبهم في هذه الناحية له أهمية كبيرة. كذلك يكون لكسب تأييد المختارين وقادة الجمهور المحليين ورجال الدين اهمية كبيرة في كسب ثقة الجمهور . كل هذه الوسائل لها قيمتها ويجب استغلالها إلى أقصى حد ممكن . ولكن مع ذلك يجب الالتجاء إلى طرق الدعاية الشعبية لاهميتها في تعريف أفراد الجمهور باغراض الاستقصاء وأهدافه وفي أعدادهم لاستقبال هذا العمل بروح ودية .

والهدف الاول لحملة الدعاية يجب أن يكون هو تعريف الناس باغراض التعداد فنحاول بكل الطرق الممكنة ان نجعلهم يدركون الفوائد التي سوف نستخدم في تحقيقها المعلومات التي سوف يعطونها ، وبمعنى آخر كيف يكون لذلك أهمية كبيرة في وضع سياسة الحكومة نحو المسائل الاجتماعية

والاقتصادية وكيف تؤدي هذه السياسية إلى رفع مستوى المعيشة في الدولة إذا كان هذا هو الهدف .

وإذا ما استطعنا ان نفسر أغراض التعداد أمكننا أن ننتقل إلى الخطوة التالية وهي توضيح طريقة اجرائه . فيجب أولا توضيح مسئولية الجمهور في اعطاء المعلومات المطلوبة (القانون) والوسيلة التي سوف تلجأ اليها الحكومة لجمعها ، ولا يجب أن يقتصر التوضيح على عملية العد فقط بزعم انها هي التي تتصل بالجمهور اتصالاً مرتبطاً وثيقاً ، ولكن يجب أن يشتمل التوضيح على الخطوات الاخرى التي سوف تمر بها المعلومات حتى يتم نشرها .

ولا يجب أن تتركز الدعاية حول الناحية السلبية فقط أي عدم استخدام المعلومات فيما يضر بصالح الأفراد ، بل يجب أن نحـــاول توضيح الجانب الايجابي كذلك بجعل أفراد الجمهور يدركون ان المعلومات تجمع لصالحهم . وفي معالجة هذه الناحية لا يجب أن نصوّر للناس آمالاً نعلم انها كاذبة حيث أن ذلك يسيء اليهم وإلينا في الأجل الطويل ، فقد ننجح في ذلـــك حالياً ولكن لا بد وأن يتولد لدى الجمهور الشعور بعدم الثقة ، الأمر الذي يجعل من الصعب أن نكسب تعاونه في المستقبل .

وللدعاية الشعبية نوعان رئيسيان من الوسائل هما الصحافة والاذاعة متضمنة التليفزيون كذلك . ويمكن الاعتماد اساساً على المقـالات في الصحف اليومية لشرح جميع الوسائل المتعلقة بالدراسة . ولكن نظراً لقلة انتشار الصحف في الدول العربية يكون الراديو والتليفزيون هما اكثر الوسائل فعـــالية خاصة بالنسبة للذين يقطنون في المناطق الزراعية .

والنوع الشائع من الدعاية بواسطة الاذاعة هو سلسلة من الاحـــاديث عن الدراسة . ولا بد ان تكون هذه الاحاديث قصيرة وأن تحتوي على ما يثير اهتمام الجمهور . كذلك يمكن تنظيم برامج يُسأل فيها احـــد المسؤولين في الجهاز القائم بالدراسة أسئلة يوجهها أفراد يمثلون طوائف الجمهور المختلفة بحيث

تكون الإجابة عليها توضيحاً للنواحي المختلفة التي يقوم عليها التعداد . كذلك يمكن تنظيم تمثيليات بسيطة توضحفيها المسائل العملية المختلفة المتعلقة بالتعداد. ويجب ان نـــلاحظ ان تنظيم برامج الدعاية بالراديو او التليفزيون هو من اختصاص الفنيين ولهذا يجب ان يترك العمل في تعداد واسع النطاق الى قسم الدعاية الذي يقوم بالعمل فيه موظفون أخصائيون في هذه الناحية .

وفي حالة الدراسات التي تجري بواسطة المعاينة والتي تقتصر على مناطق محدودة ، كذلك في حالة الدراسات التي تكون طائفة قليــلة من الجمهور هي المعنية بالدراسة (التعداد الصنعي مثلاً) يكون من الاسراف توجيه حملة الدعاية للسكانعامة ولهذا يمكن الاقتصار على النشرات التي ترسل الىالطوائف المعنية فقط وعلى المقالات التي تنشر في الجرائد والمجلات المتخصصة . وفي التعداد الزراعي حيث تكون وحدات التعداد مبعثرة في جميع انحاء الدولة لا يمكن الاعتماد كثيراً على الصحف الا اذا وجدت الصحف المحلية وهو ما لا يتوفر في الدول العربية . وفي هذه الحاله تكون الاذاعة بالاضافة الىالنشرات واللافتات التي تعرض في أسواق القرى أكثر وسائل الدعاية فعالية .

ومن الواضح ان توقيت الدعاية مسألة ذات أهمية كبيرة – فاذا ظهرت المقالات في الصحف واذيعت البرامج في الراديو والتلفزيون قبـــل التعداد باسابيع طويلة فان اثرها لا بد ان يضيع عند حلول موعد التعداد وبذلك لا تحقق الفائدة منها . وبذلك يحسن ان نبدأ الدعاية في شكل بسيط للغاية ثم تقوى تدريجياً حتى بدأ التعداد . وفي تعداد السكان يجب ان تصل حمـــلة الدعاية الى ذروتها خلال الايام القلائل التي تسبق يوم التعداد حتى يمكن بحلول هذا اليوم ان يكون كل فرد في الدولة قد اصبح لديه فكرة واضحة عما يجري حوله في هذا اليوم . أما في الدراسات التي يجري العد فيها على مراحل فان حملة الدعاية يمكن ان تستمر اثناء المدة الطويلة التي يجري فيها العد وخاصة في المناطق التي لم يشملها العد بعد وتلك التي يكون العد جاريا فيها .

وبالرغم من ان حملة الدعاية تهدف اساساً الى اثارة اهتمام الجمهور وكسب تعاونه عند اجراء عملية العد ، الا انه لا يجب ان نوقفها عند الانتهاء من هذه العملية ، فما دمنا قد اثرنا اهتمام الجمهور فلا بد ان نحافظ على هـــذا الاهتمام بنشر النتائج التي نتوصل اليهـــا في صيغة شعبية بسيطة ، ويساعد ذلك في ايقاظ الوعي الاحصائي لدى الجمهور ، الأمر الذي يساعدنا في اجراء الدراسات التالية في المستقبل

ويجب ان تتذكر دائماً ان الدعاية عملية باهظة التكاليف ، ولهذا يجب محاولة اقناع اصحاب الصحف على تقديم الفراغ المجاني لنشر المقالات القصيرة واخبار التعداد . واذا لم ننجح في ذلك يكون من المهم عمل ميزانية دقيقة للدعاية على اساس المفاضلة بين فعالية الوسائل المختلفة التي يمكن الالتجاء اليها ولا يمكن ان ننجح في ذلك الا اذا تركنا امر الدعاية للفنيين فيها وسوف نرى في النهاية ان مـــا ينفق على توظيفهم يبرره ما يستطيعون تحقيقه من وفر في تكاليف الدعاية من جهة ومن زيادة فعاليتهـــا من جهة اخرى . واذا لم نترك امر الدعاية للفنيين فاننا لا بد ان نسرف في الانفاق دون ان نحقق النتائج المرجوة .

سادساً – القيام بعد اختباري :

بعد الانتهاء من جميع الخطوات التمهيدية لاجراء العد يجب اجراء عد اختباري للتأكد من ان جميع العمليات التنظيمية التي اتخذت قادرة على تحقيق وظائفها. والفائدة الأولى من العد الاختباري هي اختبار الاستمارة في الميدان قبل القيام بالعد الفعلي إذ مهما دققنا في تصميمها فان الاطمئنان اليها لا يمكن أن يتحقق إلا بتجربتها في الميدان للكشف عن عيوب في صياغة الألفاظ أو في صياغه الأسئلة، اذ قد يتبين لنا من الاختبار أن بعض الألفاظ المستخدمة ليست من الوضوح الذي تصورناه في بادىء الأمر ، أو ان بعض الاسئلة قد فهمها المستجوبون فهما يختلف عن المقصود منها ، وبذلك يكون من الضروري

تغيير مثل هذه الألفاظ ومثل هذه الأسئلة اذا كان ذلك ممكناً أو تزويد العدادين بتعليمات أكثر تفصيلاً لمساعدتهم في شرح الأسئلة للمستجوبين وفي الحصول على الاجابات المطلوبة .

كذلك يساعدنا العد الاختباري في اختبار التنظيم الاداري الذي اعددناه لاجراء الدراسة ، فقد يتبين لنا مثلاً أن الاجراءات التي اتخذت لتوزيع الاستمارات فيها ضياع كبير للوقت وان تغيير بسيط في هذه الاجراءات يمكن أن يوفر جزءاً من الوقت الضائع . كذلك قد يتبين لنا أن بعض العدادين بالرغم من نجاحهم في الاختبارات النظرية لا يصلحون للقيام بالعمل في الميدان بسبب الخجل والارتباك عند مواجهة الناس . كذلك يساعدنا العد الاختباري في تحديد الوقت اللازم لكل وحدة استبيان وبذلك يكون من الممكن تقدير ساعات العمل الضرورية لاجراء عمليات العد ، واذا اخذنا في الاعتبار الوقت الذي يضيع في الانتقال من وحدة الى أخرى أمكننا أن نتحقق من عدد العدادين الذين يجب علينا توظيفهم لاجراء عملية العد . ولا شك أن كل هذه الاختبارات تساعدنا في التحقق من تقديراتنا للتكاليف والتأكد من المبالغ التي خصصناها لاجراء الدراسة .

وعند اجراء العد الاختباري يحسن اختيار الوحدات التي يحتمل أن نواجه معها أكبر قدر من المشاكل ، وفي نفس الوقت يحسن أن نجني بعض الخبرة في معالجة المشاكل التي يمكن أن نواجهها في مختلف أنواع الحالات . ولهذا لا يجب أن نعتبر العد الاختباري عملية صغيرة تجري في منطقة واحدة نختارها لهذا الغرض ولكن كعملية واسعة النطاق تجري في أماكن تمثل مناطق الدولة المختلفة .

ونطاق العد الاختباري يتوقف بلا شك على نوع الدراسة التي نريد القيام بها ، فمثلاً في تعداد صناعى يمكن أن يكون العد الاختباري شاملاً لعدد من المؤسسات في كل صناعة للتأكد من صلاحية الاستمارة لمختلف أنواع الصناعات

أما في دراسة ضيقة النطاق كبحث ميزانية الأسرة يمكن أن يتخذ العد الاختباري شكل اختبار دقيق للاستمارة بين بعض الأسر الشبيهة لتلك التي التي يتضمنها البحث وفي مثل هذه الدراسات ما يعنينا في الواقع هو اختبار الاستمارة وليس التنظيم الاداري الذي لا يكون له أهميته الكبيرة التي تظهر في حالة الدراسات الواسعة النطاق كتعداد السكان .

سابعاً – جمع المعلومات في الميدان

تعتبر عملية التعداد المرحلة النهائية التي تنتهي بها جميع الخطوات التمهيدية التي سبق القيام بها والتي ناقشنا أهمها فيما سبق. وفي تعداد السكان تجري عملية العد النهائية في وقت قصير جداً وهي لذلك تتضمن مجهوداً ادارياً ضخماً . وفي الأنواع الأخرى للدراسات الاستقصائية يمكن أن تجري عملية العد خلال وقت طويل ولهذا لا يكون لمشاكل التنظيم الاداري نفس التعقيد الذي نواجهه في تعداد السكان . وحتى في مثل هذه الحالات لا بد من الرقابة المحكمة على عملية العد والا استمرت طويلاً الأمر الذي يؤدي الى ضياع الوقت والمال .

واذا ما تم تجنيد العدادين وتدريبهم ، وإذا ما تم تحديد مناطق التعداد وتنظيم قوائم بوحدات العد التي تتبع كل عداد ، تكون المشاكل التي نواجهها بعد ذلك أثناء اجراء عملية العد هي المواصلات والرقابة على العمل .

المواصلات :

يجب نقل كل تعداد الى منطقة اختصاصه بحيث يتواجد فيها عند بدء عمل كل يوم ، كذلك يجب اتخاذ الاجراءات اللازمة لتأمين انتقاله من كل وحدة الى الاخرى اذا استدعى الامر ثم انتقاله في نهاية اليوم الى منزله أو أي مكان يتفق عليه وفي المدن الكبيرة حيث تكون وحدات التعداد متجاورة يكون الانتقال بينها على الاقدام وبذلك لا تنشأ مشكلة نقل العدادين بين

وحدات التعداد.أما في التعداد الزراعي أو أي استقصاء يتضمن عد وحدات متناثرة لا بدمن اتخاذ الاجراءات اللازمة لتأمين انتقال العدادين بين الوحدات المختلفة . والوسيلة المناسبة للمواصلات تتوقف بالطبع على جغرافية المكان الذي يشمله التعداد وفي التعداد الزراعي الذي أجري في العراق استخدمت جميع وسائل المواصلات من القوارب للانتقال في المستنقعات الجنوبية والجمال والبغال التي استخدمت في المناطق الشمالية البعيدة . أما في تعداد السكان اذا استثنينا المناطق القروية البعيدة التي تكون فيها الاسر متناثرة فالغالب أن العدادين يستطيعون الانتقال بين الوحدات المختلفة بواسطة الاقدام وبذلك تقتصر مشكلة المواصلات في هذه الحالة على تأمين نقل العدادين الى أماكن اختصاصهم في الوقت المناسب .

واذا كانت وسائل المواصلات العامة منظمة تنظيماً جيداً يكون كل ما نحتاجه هو تزويد العدادين بتصاريح تسمح لهم بالسفر المجاني مستخدمين هذه الوسائل (التابعة للبلدية أو الحكومة) . أما في الأماكن التي لا تصل اليها هذه الوسائل العامة والتي يجب على العدادين زيارتها يكون واجباً على الادارة الاحصائية أن تنظم وسائل المواصلات الخاصة بها وهذا اما بتأجير العربات الضرورية أو بشرائها ثم بيعها عندما ينتهي العمل . واذا كان التعداد يستغرق وقتاً طويلاً يكون شراء العربات هو أفضل الطرق التي يمكن أن تلجأ اليها الادارة الاحصائية لتأمين وسائل المواصلات من الناحية الاقتصادية حيث أن التأجير يكون في التالب أكثر تكلفة . وقبل تأجير أو شراء أي من وسائل المواصلات يجب بحث امكانية استخدام الوسائل العامة بحثاً دقيقاً شاملاً لان ذلك يكون حتماً أقل تكلفة

مراقبة عملية التعداد :

مهما كان مستوى العدادين عالياً ومهما كان تدريبهم دقيقاً فان الاشراف والرقابة على عملية العد من الأهمية بمكان كبير . ولهذا يجب التأكد من تنظيم

مناطق التعداد تنظيما يساعد على امكانية الاشراف والرقابة عليها أثناء اجراء عملية العد . ونوع الاشراف يختلف بطبيعة الحال تبعاً لنوع الاستقصاء الذي يطلب اجراؤه . وفي استقصاء واسع النطاق مثل تعداد السكان والتعداد الزراعي قد يكون أمراً ضرورياً تنظيم سلسلة مفصلة من الرقابة تبدأ من المكتب المركزي حتى تصل الى العداد في الميدان. وعدد الحلقات التي تشملها تلك السلسلة يتوقف على نطاق الاستقصاء وعلى درجة شموله لمناطق الدول المختلفة . وفي أي استقصاء واسع النطاق على كل حال ، لا يمكن تجاهل التنظيم المحلي .

وعندما تكون الادارة المحلية منظمة تنظيما جيداً وبالأخص في حالة الدراسات الاستقصائية الواسعة النطاق مثل تعداد السكان يحسن أن يوكل أمر الادارة الى السلطات المحلية تحت الاشراف الفني للادارة الاحصائية . وتعداد السكان يؤثر أحياناً على القانون والنظام في بعض المناطق ولذلك يحسن منح المسؤولية بطريقة يمكن بواسطتها تلافي أي ارتباك في الادارة المحلية . وحتى لو لم تجعل الادارة المحلية مسؤلة مباشرة عن اجراء التعداد يكون من المهم جداً اجراء التنظيم بالتعاون معها اذ يمكن أن يكون للاحتكاك معها آثار وخيمة بالنسبة لنجاح التعداد بينما يكون للتعاون الصادق معها أثره الحميد في كسب تعاون الجمهور وتجنب الصعوبات السياسية .

وعدد المشرفين الذين نكون في حاجة اليهم يتحدد تبعاً لعدد العدادين الذين يمكن للمشرف الواحد الاشراف عليهم . والمشرفون على العمل في الميدان يعلنون تقاريرهم الى المسئولين المحليين ، فإذا كان عدد المناطق ليس كبيراً يمكن لهؤلاء أن يبلغوا تقاريرهم إلى الادارة المركزية وبذلك يكون عملهم حلقة اتصال بين هذه الادارة وبين القائمين بالعمل في الميدان . وإذا كانت السلطة المحلية هي المسئولة عن اجراء التعداد يكون من الواجب تحديد شخص واحد في كل ادارة يكون مسئولاً عن الاتصال بالادارة المركزية للتعداد .

ويجب على كل مشرف أن يكون قادراً على الاتصال المتواصل بالعدادين أثناء قيامهم بالعمل في الميدان ، كما يجب تنظيم الاجراءات التي يستطيع بمقتضاها العدادون الاتصال بالمشرفين عليهم عندما يواجهون صعوبات تحتاج إلى العلاج السريع . وبذلك تكون أول وظيفة للمشرف هي أن يكون موجوداً في مكان معين ومستعداً للاجابة على الأسئلة المعقدة التي يمكن أن تظهر أثناء العمل ولمساعدة العدادين في مواجهة المستجوبين الذين يمتنعوا عن اعطاء المعلومات المطلوبة .

والوظيفة الثانية للمشرفين هي التأكد من قيام العدادين بعملهم على الوجه الأكمل . ففي بعض الحالات عرف ان العدادين يقومون بملأ الاستمارات بأنفسهم وهم جالسون في المقاهي ، والاشراف الدقيق يجب أن يعمل على منع هذا من الحدوث ، وان كان التدقيق في اختيار العدادين وتدريبهم التدريب الحسن هو أفضل طريقة لتلافي أي تراخي يمكن أن يحدث .

والوظيفة الثالثة للمشرفين هي مراقبة تقدم العمل حتى يمكن التأكد من اتمامه في الوقت المحدد . فقد يظهر ان الوقت المحدد لاجراء التعداد في منطقة ما كان أكثر من اللازم بينما الوقت الذي حدد لمنطقة اخرى أقل من اللازم وفي مثل هذه الحالة يصبح من الضروري تحويل بعض العدادين من المنطقة الأولى الى الثانية وان كان يجب أن نتجنب مثل هذا الوضع بكل ما يمكن من الوسائل .

ولمساعدة المشرفين في هذه الوظيفة يجب على كل عداد أن يسجل ما قام به من عمل أولاً بأول ، ويمكن اجراء ذلك بتزويد العداد بقائمة بالوحدات التي يتحتم عليه زيارتها ويطلب منه ان يؤشر على كل وحدة ينتهي من جمع المعلومات عنها مع بيان الملاحظات التي تعن له بخصوصها . وفي نهاية اليوم تجمع هذه التقارير وتسلم للمشرفين حتى يمكنهم مراقبة تقدم العمل ، وبذلك يمكن التعرف على العدادين الذين يتسم عملهم بالبطء والذين يواجهون حالات كبيرة من عدم الاجابة فيوجه اليهم رقابة أكثر حتى يمكن استبدالهم إذا لزم

الأمر . فهناك أشخاص لا يمكن أن يكونوا عدادين أكفاء حيث لا يستطيعون تجنب معاداة المستجوبين ، وهذه الحقيقه قد لا تبدو واضحة إلا عند مـــا ينزلون إلى الميدان حيث تظهر عدم كفاءتهم في كثرة عدد حالات عدم الاجابة التي يبلغونها إلى المشرفين عليهم . ومثل هؤلاء العدادون يجب تحويلهم إلى أعمال اخرى بأسرع ما يمكن .

وفي نهاية كل يوم يجب تسليم الاستمارات التي تم ملأها الى المشرفين الذين يقومون بمراجعتها مع التقارير المقدمة من العدادين . فاذا وجد أن عدد الاستمارات التي انتهى منها أقـــل بكثير من العدد المقدر في بادىء الأمر ، يكون ذلك دليلاً على عيب في التنظيم العام للاستقصاء أو في الاستمارة لم تتنبه اليه هيئة التعداد . وفي هذه الحالة يجب أن تخطر الادارة المركزية بالأمر حتى يمكن بحث أسباب هذا العيب وعـــلاجه إذا أمكن ، وفي الدراسات الاستقصائية التي تستغرق عملية العد فيها وقتاً طويلاً يكون من الضروري مراقبة اجراء المراحل الأولي من العمل رقابة محكمة حيث يكون هناك متسع من الوقت لعلاج أي عيب في التنظيم أو في الاستمارة بشرط التنبه إلى هذه العيوب في وقت مبكر . ولا يكون هذا ممكناً في الدراسات التي تتركز فيها عملية العد في وقت قصير وبذلك ينحصر جهد المشرف في التأكد من شمول العد لجميع الوحدات في الوقت المحدد إذ لا يكون هناك وقت لعلاج أي عيب يظهر في التخطيط الأصلي .

ويجب مراجعة الاستمارات المنتهية مراجعة سريعة في نهاية كل يوم للتأكد من عدم وجود إجابات غير منطقية على بعض الأسئلة ومن الاجابة على جميع الأسئلة . وهذه المراجعة المبدأية في الميدان ذات أهمية كبيرة حيث يكون من السهل اتخاذ الاجراءات اللازمة لزيارة المستجوبين مرة ثانية إذا وجد في اجاباتهم أخطاء واضحة ، بينما لا يكون ذلك سهلا إذا وصلت الاستمارات إلى الادارة المركزية دون التنبه الى مثل هذه الأخطاء في الميـــدان . كما أن معلومات العاملين في الميدان المحلي عن المستجوبين لها فائدة كبيرة في المساعدة في

اجراء مثل هذه المراجعة والتي لذلك يجب أن تعتبر جزءاً هاماً من أعمال الميدان .

ثانياً – تدقيق المعلومات

ان عملية مراجعة البيانات التي جمعت في الاستمارات للتأكـد من صحتها قبل اجراء التبويب عليها تسمى بعملية تدقيق الاستمارات ، وتهدف هـــذه العملية الى تحقيق الامور الآتية :

١ – التأكد من ان البيانات التي جمعت صحيحة ودقيقـــة الى اكبر درجة ممكنة .

٢ – التأكد من تماسك البيانات فيما بينها وعدم تعارضها مع الواقع .

٣ – التأكد من تسجيل البيانات في الاستمارات تبعاً للتعليمات الخاصة بذلك

٤ – التأكد من شمول الاستمارات لجميع البيانات المطلوبة .

٥ – التأكد من تسجيل البيانات تسجيلاً يساعد على التبويب .

٦ – التأكد من تنظيم البيانات في الاستمارة تنظيما يساعد على ترميزهـــا وتبويبهـــا .

٧ – ملاحظة تعليقات العدادين والمشرفين عليهم حيث يمكن أن يكون لهذه التعليقات فائدتها في تفسير نتائج الدراسة .

ويجب أن نشير قبل مناقشة عملية تدقيق الاستمارات الى ان جمع المعلومات يجب أن ينظم بحيث لا يكون هناك حاجة إلى تصحيح أو تغير في البيانات التي جمعت الا في الحالات النادرة فقط ، فاقتصادا الوقت يكون من الأفضل منع الخطأ بدلاً من وقوعه ثم محاولة تصحيحه . وقد سبق أن ناقشنا الوسائل المختلفة التي يمكن بواسطتها منع الخطأ قبل اجراء الدراسة الاختبارية وتجنيد

العدادين تبعاً للمستوى المطلوب وتدريبهم التدريب الكافي والاشراف عليهم في الميدان .. الخ ..

خطوات تدقيق الاستمارات

في الدراسات الواسعة النطاق تبدأ عملية تدقيق الاستمارات في الميدان حيث يجب أن يطالب العدادين بمراجعة الاستمارات التي ينتهون منها قبل تسليمها للمشرفين عليهم ، ويحسن أن تتم هذه الخطوة في مكان المستجوب وقبل مغادرته ، وعند استلام المشرفين الاستمارات يجب أن يتأكدوا من الآتي :

١ – وجود اجابة لكل سؤال في الاستمارة ، حتى ولو كانت هذه الاجابة مجرد تسجيل علامة معينة .

٢ – كتابة الاجابات بخط واضح يمكن قراءته دون خطأ .

٣ – ان الاجابات سجلت في أماكنها طبقاً للتعليمات الخاصة بها .

٤ – التماسك بين اجابات الأسئلة المختلفة ، بمعنى التأكد من عدم وجود تعارض بين العمر والحالة الزواجية مثلاً ، بين المهنة والحالة التعليمية ، بين عمر الام وعمر الابن البكر ، بين عدد العمال في المؤسسة وانتاجها ، بين مجموع مساحات الحيازة المستخدمة في الاغراض المختلفة والمساحة العامة للحيازة ، بين دخل الأسرة ومجموع انفاقها اذا اشتملت الاستمارة على اسئلة تبين الادخار والاقتراض .. الخ .

٥ – إذا ظهر عدم تماسك بين اجابات بعض الأسئلة ، هل تنبه العداد إلى ذلك وهل سجل في الملاحظات تفسيراً لهذا التعارض .

٦ – ان البيانات المسجلة تعكس بدقة اجابات المستجوبين .

٧ – عدم وجود أي خطأ في أسماء وعناوين المستجوبين كما هي مسجلة في الاستمارات .

٨ – وجود امضاء العداد وتاريخ الانتهاء من الاستمارات .

٩ – عدد الاستمارات في حوزة كل عداد وعدد الذي سلمه منها .

ويجب ان تجري عملية المراجعة من قبل المشرفين أولاً بأول ، أي يجب عليهم مراجعة كل ما تسلموه من الاستمارات دون انتظار انتهاء عملية جمع المعلومات ، إذ بذلك يمكن تنبيه العدادين الى أخطائهم حتى لا تتكرر معهم في الاستمارات التالية . كما ان هذه المراجعة يمكن أن تنبه المشرفين الى سوء فهم العدادين لبعض الأسئلة أو التعليمات الخاصة بها أو لكيفية توجيه الأسئلة أو لكيفية تسجيلها في الاستمارات وبذلك يمكل تصحيح ما فهموه خطأ أثناء تدريبهم كذلك فان المراجعة بهذة الطريقة يمكن أن تساعد العدادين على تذكر اجابات المستجوبين نظراً لعدم مضي وقت طويل على مقابلتهم وبذلك يمكن تصحيح بعض الأخطاء دون الحاجة إلى إعادة المقابلة مرة ثانية ، الأمر الذي يمكن أن يضايق المستجوبين كثيراً . كذلك فان المراجعة بهذه الطريقة يمكن ان تنبه المشرفين الى العدادين الذين يستمرون في تسليم استمارات فيها نفس الأخطاء التي وقعوا فيها أولاً وبذلك يمكن استبعادهم من أعمال الميدان واحلال آخرين محلهم .

كذلك فان المراجعة بهذه الطريقة يمكن أن تكشف عن وجود نقص في التعليمات فيما يختص بتفسير بعض الأسئلة أو بعض المصطلحات المستخدمة أو بعض التعاريف ، الأمر الذي يساعد على معالجة هذا النقص في وقت مبكر قبل أن تنتهي أعمال الميدان ويصبح من المستحيل اجراء العد مرة أخرى .

والخطوة التالية هي مراجعة الاستمارات في المكتب القائم بالدراسة مراجعة تفصيلية ، ذلك لأن مراجعة الميدان نظراً للسرعة التي تجري بها تكون

مراجعه عامة لا يمكن الاعتماد عليها كلية ، أو تقتصر فائدتها على اكتشاف السهو أو الأخطاء الواضحة البسيطة حتى يمكن تصحيحها في الميدان عندما تكون الاستمارات لا زالت قريبة من المستجوبين . ويكون الهدف من المراجعة في المكتب التأكد من ان البيانات التي جمعت ليس بها أي تعارض يمكن ان يكشف عن أخطاء نتجت عن أي سبب من الأسباب .

وتنشأ مشكلة إدارية عند إجراء المراجعة المكتبية ، وهي هل يقوم بمراجعة الأستمارة الوحدة مراجع واحد ، أو هل من الأفضل تقسيم الاستمارة الى أقسام يقوم بمراجعة كل منها مراجع خاص ؟ ولا شك ان الطريقة الثانية تكون هي الأفضل اذا كانت استمارة البحث طويلة ومعقدة حيث انها تساعد المراجع على التخصص في ناحية معينة يستطيع أن يجيدها بالبحث والمران . إلا ان التوسع في تقسيم عملية المراجعة يمكن ان يكون خطراً حيث انه لن يساعد على الكشف عن عدم التماسك بين البيانات في الأقسام المختلفة ، الأمر الذي لا يمكن ان يكشف إلا بالمقارنة بين إجابات الأسئلة المختلفة في الاستمارة . على انه يمكن الجمع بين الطريقتين في المراجعة وذلك بتكليف مراجع واحد مراجعة الاستمارة بأكملها من حيث التماسك بينما يقوم مراجعون آخرون متخصصون بمراجعة بعض الأرقام في الاستمارة التي تحتاج الى مراجعة خاصة . ومن المستحسن اذا كانت الاستمارة تحتوي على بيانات تحتاج مراجعتها الى إجراء عمليات حساب ان تكون من اختصاص مراجع مزود بآلة حاسبة تحقيقاً للسرعة والدقة في العمل .

ولكي يستطيع المراجع تدقيق أي بيان لا بد أن يكون لديه علم كافي بالموضوع حتى يستطيع أن يتبين الاجابات غير المحتملة أو المفتعلة أي التي يكتبها العدادون في الاستمارات دون مقابلة المستجوبين المعنيين . ويكون هذا العيب بالغ الخطورة في الدراسات المتعلقة باراء المستجوبين حيث لا تكون البيانات مصورة لآرائهم بل لآراء العدادين، الأمر الذي يحكم على الدراسة بالفشل الذريع وتلافياً لهذا العيب يمكن مطالبة العدادين بتدوين تعليقات

المستجوبين وملاحظاتهم ، ولا شك أن العدادين لن يكونوا من الذكاء والمهارة بحيث يستطيعوا تصور ملاحظات وتعليقات تختلف فيما بينها اختلافاً جوهرياً وبحيث تبدو مقبولة من المراجعين . وتزداد هذه المشكلة تعقيداً اذا قابل العدادون مستجوبين غير الذين حددوا لهم حيث تأتي تعليقاتهم وملاحظاتهم متباينة ومقبولة ، الأمر الذي لا يمكن أن يكشفه المراجعون . والوسيلة الوحيدة لاكتشاف مثل هذه الاجابات هي التأكد من عمل العدادين في الميدان ثم اجراء اختبار على الاستمارات بارسال مراجعين لمقابلة عينة من المستجوبين مرة ثانية .

وفي بعض الأحيان يظهر عدم دقة الاجابات من مراجعة البيان الخاص بمجموعة من المستجوبين وليس بواحد منهم فقط ، ويكون ذلك بمقارنة المجموع الخاص ببيان ما بما نتوقعه ، فاذا لاحظنا فرقاً كبيراً كان ذلك دليلاً على أن المستجوبين لم يتحروا الدقة عند الاجابة عن السؤال الخاص بهذا البيان . ومثلاً على ذلك ما حدث في تعداد السكان عام ١٩٠٠ في أمريكا حيث تبين بعد التبويب أن عدد الزنوج الذين لا يستطيعون الكلام باللغة الانجليزية كبير إلى درجة توحي بالشك . وقد ظهر من تدقيق الاستمارات أن ترتيب الأعمدة الخاصة بالأسئلة – القدرة على القراءة باللغة الانجليزية والقدرة على الكتابة بها والقدرة على التحدث – أدى إلى هذا الخطأ ، حيث كانت الأعمدة متلاصقة فكان العداد نتيجة للسرعة يكتب لا ، لا ، لا . في الوقت الذي كان يجب أن يكتب فيه لا ، لا ، نعم ويتضح من هذا المثال أن هذا الخطأ لم يكن من الممكن أن يكتشف بمراجعة استمارة كل مستجوب على حدة ، أما بمراجعة الاستمارات ككل اتضح الخطأ ، الأمر الذي ساعد على تصميم الاستمارة تصميماً مختلفاً عند اجراء التعداد التالي .

تماسك الاجابات

ان من أهم عناصر عملية تدقيق الاستمارات التأكد من عدم وجود تعارض

بين اجابات الاسئلة المختلفة ويكون ذلك باجراء المقارنة بين اجابات الأسئلة المرتبطة ببعضها طبقاً لنظام منصوص عليه في التعليمات الخاصة بالمراجعين . وفي بعض الأحيان يكون من الممكن تحديد الخطأ في اجابة أي من الأسئلة التي ظهر التعارض فيما بين اجاباتها وبذلك يمكن اجراء التصحيح في المكتب دون الرجوع إلى المستجوب . وفي حالات أخرى لا يمكن التحديد وبذلك يجب الرجوع إلى المستجوب لاجراء التصحيح وإذا كان مستحيلاً يجب تصنيف اجابة هذا المستجوب ضمن فئة « لم يبين » .

وغالباً تكون عملية مقارنة الأرقام التي تتضمنها الاجابات المختلفة عملية منفصلة يقوم بها مراجع خاص تكون وظيفته التأكد من ان المجاميع تتطابق مع الأجزاء التي تكونها ، وان كل رقم يقع ضمن نطـــاق المنطقة الخاص بالبيان المتعلق به ، وان كل رقم لا يتعارض مع أي رقم آخر في الاستمارة ، وان الأرقام التي تظهرفي استمارات العدادين لا تختلف اختلافاً كبيراً عن الارقام الموجودة في استمارات العدادين الآخرين (في الدراسات الخاصة بطبقة معينة من المجتمع) . فاذا تبين وجود تعـــارض في أي حالة من هذه الحالات لا بد من التحقق من الأرقام أو من وجود تفسير بذلك عند العدادين سجلوه في ملاحظاتهم على ما جمعوه من بيانات .

وحدة القياس :

بالرغم من أن الاجابات قد تكون سليمة الا انه لا يمكن تبويبها إلا بعد تحويلها إلى وحدات قياس موحدة . فإذا كان العدادون قد لاحظوا التعليمات الخاصة بذلك عند جمع البيانات فان البيانات سوف تكون مسجلة على أساس موحد ، ولكن مع ذلك يجب اجراء المراجعـــة التأكيد من ان الامر هو هكذا . كذلك فان بعض المستجوبين يمكن أن يكونوا قد اجابوا بتفصيل غير مطلوب فتكون وظيفة المراجعين شطب التفاصيل حتى لا يختلط الامر على المرمزين مثلا اجاب المستجوب على السؤال الخاص بالعمر فاعطى سنه

بالسنوات والشهور والايام ، او أجاب عن السؤال الخاص بدخله فأعطى رقماً باللبرات والقروش وسجل العداد الاجابة كما لفظها المستجوب . كذلك يجب التأكد من أن طريقة تسجيل الاجابات موحدة بين الاستمارات المختلفة ، مثلا استعمال العلامة – لتدل على عدم انطباق السؤال على المستجوب بالنسبة لبعض العدادين واستعمالها لتدل على انطباق السؤال لعدادين آخرين .

وعموماً يمكننا أن نلاحظ أن وحدة القياس التي جمعت وسجلت تبعاً لها البيانات في الاستمارات المختلفة يمكن أن تتحقق إذا كانت التعليمات الخاصة بها واضحة وإذا تمسك العدادون بهذه التعليمات عند العمل في الميدان ولاحظها المشرفون عند اشرافهم على العدادين .

السهو :

يمكن أن تأتي الاستمارات الى المكتب وبها بعض السهو أي بها اسئلة لا يظهر أمامها ما يدل على ان العداد قد وجهها الى المستجوبين وحصل على اجابة خاصة بها ، ولعل ذلك هو السبب – كما أشرت سابقاً – إلى أن التعليمات يجب أن تنص على تسجيل اجابة عن كل سؤال في الاستمارة ويمكن أن تكون الاجابة مجرد تسجيل علامة معينة . كذلك يمكن أن يكون السهو فيما يختص بتسجيل جزء من الاجابة ، مثلاً عدم وضع علامة الليرة بجانب الرقم الخاص بالدخل . وعلامة الفدان بجانب الرقم الخاص بالمساحة المزروعة بأي محصول من المحاصيل ... الخ ... مثل هذه الاستمارات سوف تهمل عند اجراء الترميز ، الأمر الذي يضخم فئة لم تبين الى درجة كبيرة والذي يجعل المطلعون على نتائج الدراسة لا يعولون كثيراً عليها . لذلك يكون واجب المراجعين التأكد من عدم وجود أي سهو في الاستمارات ، وكثير من الاجابات الناقصة يمكن معرفتها في المكتب باستنتاجها من اجابات الأسئلة الاخرى مثلاً السؤال هل تملك سيارة ؟ نعم – لا ، ثم السؤال ما هي ماركة السيارة ؟ فاذا وجدت اجابة خاصة بماركة السيارة أمكن معرفة أن اجابة

السؤال الأول هي نعم. على أن بعض الاجابات غير المسجلة يكون من الصعب تحديدها في المكتب ولمـا كان من المستحيل في بعض الاحيان الرجوع الى المستجوبين لذلك يجب مطالبة المشرفين في الميدان بالتنبه الى هذا العيب في الاستمارات عند جمعها من العدادين ، وبذلك عندما ترد الاستمارات الى المكتب لا يجب أن يظهر فيها هذا العيب الا نادراً .

كذلك يجب أن يتأكد المراجعون من عدد الاستمارات ، والواقع أن هذه العملية يحسن أن تكون من اختصاص موظف خاص يقوم بعد الاستمارات عند استلامها من الميدان ، إلا ان ذلك لا يمنع المراجعين من اعـادة عد الاستمارات بعد الانتهاء من مراجعتها وقبل تسليمها لقسم الترميز حيث يمكن أن يختفي بعض منها في قسم المراجعة نتيجة سهو احد المراجعين أو بعض منهم. كذلك يجب على المراجعين التأكد من استبعاد أي نسخة من الاستمارات حيث ان استمرار وجود هذه النسخ عند الترميز والتبويب فيما بعد يؤدي الى أخطاء جسيمة في النتائج ، وتوجد هذه النسخ أحياناً عند مطالبة العدادين باعادة كتابة استمارة ما أو بعض الاستمارات بسبب كثرة الاخطاء أو سبب عدم وضوح الخط المكتوب به البيانات وتسليم النسخ مع اصولهـا لقسم المراجعة مرة ثانية .

قبول أو رفض الاستمارة :

من وظائف المراجعين أن يقرروا ما إذا كانت استمارة ما دقيقة وكاملة بحيث يمكن قبولها للتبويب او انها ليست كذلك فيجب رفضها . وتنشأ هذه المشكلة بشكل واضـح عند الدراسة بالمعاينة حيث يجب التأكد من ان الاستمارات تحتوي على البيانات الخاصة بالوحدات التي سحبت في العينة وليس بوحدات اخرى . ولا شك ان اكتشاف هذا الخطأ يكون من وظيفة قسم مراقبة عملية المعاينة ، إلا ان المراجعين يجب أن يتأكدوا من عدم وجود هذا الخطأ قبل اجراء التبويب .

ويحسن أن تبين التعليمات إلى المراجعين المستوى الذي يمكن على أساسه قبول أو رفض الاستمارة وعلى المراجعين أن يتقيدوا بهذا المستوى ولا يحيدوا عنه . على أن رفض بعض الاستمارات يؤثر كثيراً على نتائج الدراسة ولهذا يحسن أن يعاد النظر في أمر هذه الاستمارات من قبل رئيس قسم المراجعين والخبير في المعاينة (اذا كانت الدراسة بالمعاينة) .

اعادة تنظيم بنود الاستمارة :

لا يكون ترتيب الاسئـــلة في الاستمارة ، احيانا ، متفقاً مع الترتيب الضروري لتسهيل عملية الترميز والتبويب ، لهذا تنقل البيانات إلى استمارات اخرى بالترتيب المطلوب لقسم التبويب . على أن هذه الخطوة يحسن تلافيها نظراً لاضاعتها للوقت ولاحتمال الخطأ عند النقل . ويمكن الاستعاضة عـن ذلك باعادة ترقيم الاسئلة بلون واضح وبارز .

التعليمات الى المراجعين :

يجب أن تطبع التعليمات الى المراجعين في كتيب خاص يكون جاهزاً قبل اجراء العمل في الميدان . ويحسن أن يرفق بهذا الكتيب نماذج من استمارات فيها امثلة من الاخطاء والعيوب التي يحتمل أن تظهر لتكون مرشداً عملياً لاجراء عملية المراجعة . ويمكن أن يحتوي كتيب التعليمات على الآتي :

١ ـ على المراجعين ان يكونوا ملمين إلماماً كاملاً بالتعليمات الموجهة الى العدادين والمرمزين بجانب المامهم بالتعليمات الخاصة بهم حيث ان ذلك يساعدهم على اكتشاف الاخطاء المختلفة .

٢ ـ لا يجب على المراجعين بأي حال من الأحوال ان يقضوا بالمسح او الشطب على معالم البيان الاصلي المسجل في الاستمارة عند تصحيحه .

٣ ـ يجب استخدام قلم من لون آخر عند إجراء التصحيح .

٤ ــ لا يجب إجراء أي تصحيح إلا إذا وجد ما يبرر ذلك من إجابات الأسئلة الأخرى في الاستمارة .

٥ ــ يجب تحديد البيانات التي يمكن للمراجعين تغييرها والبيانات التي لا يجب عليهم تصحيحها إلا بعد أخذ رأي رئيس القسم .

٦ ــ يجب استشارة رئيس القسم فيما يختص بالاستمارات للتي يرى المراجعون استبعادها . ولا يجب وضـــع هذه الاستمارات في الملف الخاص بها الا بعد أخذ رأي المشرف على عمليه المعاينة (في حالة الدراسات بالمعاينة).

٧ ــ يجب وضع العلامة الخاصة بالمراجع عند تصحيح أي بيان حتى يتمكن من يراجع بعده أن يعرف سبب هذا التصحيح .

٨ ــ يجب وضع التعليمات الخاصة بكيفية اجراء التصحيـح او التغيير ، مثلاً شطب البيان بخط أفقي لإلغاء هذا البيان إلغاء تاماً والشطب بخطوط رأسية إذا أريد إلغاء جزء من البيان وشطب البيان وإجراء التصحيح فوقه أو تحته في مكان واضح إذا أريد التصحيح .

٩ ــ يجب تنبيه المراجعين الى ان عملهم يمكن ان يراجع مرة ثانية ولذا يجب أن يكـــونوا مستعدين لتفسير أي تغيير أو تصحيح أجروه على الاستمارات .

١٠ ــ يجب اثبات التاريخ على الاستمارات المنتهية من المراجعة وكذلك توقيع المراجع حتى يمكن مراقبة عمل المراجعين من ناحية السرعة والدقة .

كذلك يجب أن يشتمل كتيب التعليمات على تفسير خاص بكل بيان في الاستمارة ، كيف يكون البيان مقبولاً وكيف لا يكون كذلك إلا اذا كان مصحوبا بتفسير خاص . كذلك يجب ان توضح التعليمات أجزاء الاستمارة التي يراجع تماسكها مع الأجزاء الأخرى . وحيث ان البيانات تكون مرتبة في الاستمارة لتسهيل عملية جمع المعلومات في الميدان ، لذا يكون من المرغوب فيه

أن توضح التعليمات افضل ترتيب للمعلومات عند إجراء المراجعة إذا كان هذا الترتيب يختلف عن الترتيب الموجود في الاستمارة . كذلك يحسن ان تُظهر التعليمات البيـــانات التي يحتمل ان تظهر فيها الاخطاء وكيفية اكتشافها عند المراجعة . ومن الأهمية بمكان كبير أن تبين التعليمات الأسس التي بناء عليها يمكن رفض الاستمارات أو قبولها لإجراء عمليات التبويب عليها ، وتتعلق هذه الأسس غالباً بشمول الاستمارة للبيانات المطلوبـــة وصحة هذه البيانات ودقتها ، وكذلك بصحة اعتبار الاستمارة ممثلة لوحدة في العينـــة اذا كانت الدراسة تجري بالمعاينة .

كذلك يجب أن يساعد تنظيم عملية المراجعة على حصول كل مراجع على نسخة من أي تعديل أو تغيير في التعليمات الخاصة بهذه العملية . إلا انه يجب أن نلاحظ ان التعليمات الخاصة بعملية المراجعة لا بد من إعدادها بحيث لا يجري عليها أي تغيير إلا في الحالات القصوى فقط حيث ان تعديل التعليمات من حين لآخر يؤدي الى ضياع الوقت ، إذ يترتب على هذا التعديل إعادة مراجعة الاستمارات التي انتهت مراجعتها ، كما ان المراجعين يفقدون ثقتهم واحترامهم للمشرفين عليهم ، فضلاً عن فقدان حماسهم للعمل وتقاعسهم عن الاستفسار إذا واجهتهم اية مشكلة ، وبذلك يفقد المشرفـــون على المراجعة سلطتهم وقدرتهم على ضبط العمل .

ومن الأهمية بمكان كبير ان يفهم المراجعون ان لهم الحق في الاستفسار إذا واجهتهم اية مشكلة فنية ، ولا بد أن يكون هناك تنظيم خاص بذلك حيث يمكن ان تكتب هذه الاستشارات على بطاقات مقاس ٣ × ٥ بوصة وعند ورودها إلى رئيس قسم المراجعة تصنف في ملفات خاصة حسب نوع البيان أو حسب أقسام الاستمارة مع الاجابات والتوضيحات الخاصة بها، وبذلك يمكن الرجوع اليها قبل توضيح أي استفسار جديد حتى لا يحدث أي تعارض فيما بينها . وإذا قدم الاستفسار من عدة مراجعين يحسن أن يصدر رئيس

القسم نشرة تفسيرية كي توزع على الجميع . ولا بد من الاحتفاظ بهذه الملفات كي تسلم بعد إنهاء عملية المراجعين إلى القائمين بكتابة التقرير على الدراسة حتى تكون مرشداً لهم عند تحليل البيانات والتعليق عليها .

تنظيم قسم المراجعة :

في الدراسات الضيقة النطاق يمكن ان يقوم مراجعان بالعمل ،ومن الخطأ أن يعتقد المشرف على الدراسة انه يستطيع أن يشرف على جمع المعلومات وفي نفس الوقت يقوم بالمراجعة حيث انه عند جمع المعلومات تنشأ مشاكل كثيرة وملحة لا تترك وقتاً كافياً للمشرف كي يقوم بعملية المراجعة في الوقت المناسب .

وفي الدراسات الواسعة النطاق يجب ان يكون هناك قسم خاص للقيام بعملية المراجعة ويجب أن يعمل هذا القسم على أساس يمكن به ضبط العمل من ناحية السرعة والدقة ، ولتحقيق ذلك يجب أن يحدد متوسط عدد الاستمارات التي يمكن مراجعتها في وحدة زمن معينة ، كما يجب أن يميز عمل كل مراجع حتى يمكن تحديد المراجعين المتراخين في العمل والذين لا يتمشى عملهم مع مستوى الدقة المطلوب .

وغالباً ينشأ الاحتكاك بين المراجعين والعدادين ، الأمر الذي يوقع المشرفين على عملية المراجعة في مشاكل كثيرة . وينشأ هذا الاحتكاك لأسباب كثيرة منها ان العدادين يعتقدون ان مطالبة المراجعين لهم بتوخي الدقة التامة في تسجيل البيانات ليس إلا أمراً نظريا ومظهر من مظاهر محاولة المراجعين إظهار سلطتهم . كذلك يعتقد العدادون أن عملهم في الميدان أكثر صعوبة من عمل المراجعين الذين يجلسون في مكاتبهم ثم يرفضون استمارة ما لمجرد ان بيان ما لم يسجل فيها . وعلى رئيس قسم المراجعة ان يمنع هذا الاحتكاك من التوسع حتى لا يكون ذلك سبباً من أسباب تعطيل العمل . ولهذا يجب أن

يدرب المراجعين على كتابة ملاحظاتهم بشكل موضوعي دون المساس بشخص العداد نفسه ، وإحدى الوسائل لتحقيق موضوعية المراجعه أن يشير المراجع عند تصحيح أي بيان إلى الصفحة والفقرة من التعليمات التي تتعلق بموضوع التصحيح . كذلك يمكن تدريب العدادين على عملية المراجعة ، حيث ان هذا التدريب ولو انة ليس الا تدريباً مقتضباً يجعل العدادين يقدرون الحاجة إلى الدقة في العمل ويعطيهم فكرة أفضل عما يمكن أن تكون عليه الاستمارة مقبولة لإجراء عملية التبويب عليها فيما بعد . ويعتقد البعض ان هذا التدريب للعدادين يجعلهم قادرين على اصطناع البيانات التي تنال رضى المراجعين ، وهذا أمر خطر للغاية حيث ان الدراسة بذلك لا تصور واقع الأمر وإنما تصور ما يتخيله العدادون من بيانات . على انه في الدراسات الواسعة النطاق والتي تشتمل الاستمارة فيها على أسئلة كثيرة لا يكون لدى العدادين الوقت لاصطناع البيانات خاصة إذا كانت عملية الإشراف على العدادين والمراجعة منظمة ودقيقة .

ويجب أن يكون عدد المراجعين في المكتب كافياً لمراجعة الاستمارات بمجرد وصولها الى القسم القائم بالعمل حتى يمكن تنبيه العدادين الى أخطائهم والاستفادة منها عند تكملة العمل في الميدان . وفي بعض الحالات تطالب المكاتب المحلية بارسال بعض من الاستمارات المنتهية إلى قسم المراجعة لأخذ فكرة عن مستوى العمل في الميدان ، فاذا أتبع هذا الاجراء يجب ان ينبه على هذه المكاتب بعدم ارسال الاستمارات التي يعتقدون أنها ممتازة وأنه يجب أن تنتخب عينة عشوائية من الاستمارات بجانب الاستمارات التي يعتقدون أن بها صعوبات تحتاج لدراسة المتخصصين في عملية المراجعة .

وفي بعض الدراسات تنظم عملية المراجعة بحيث تجري محلياً في مناطق الدولة المختلفة وحتى يمكن توحيد الأسس التى تجري تبعاً لها هذه العملية يسمح لرئيس كل مكتب محلي أن يصدر التعليمات والتفسيرات للموظفين الذين يعملون معه على ألا يعمل بها إلا بعد موافقة الرئيس العام للمراجعة في المكتب المركزي .

الفصل الثالث

التصنيف والتبويب

تبين لنا من الفصل السابق أن المعلومات تجمع في استمارات كثيرة يصل عددها في بعض الدراسات الى عشرات الآلاف . والمعلومات بهذا الشكل لا يمكن أن توضح لنا اتجاهات الظواهر التي نريد دراستها ، بل لا يمكن بأي حال من الأحوال أن تساعدنا في اجراء التحليل الرياضي عليها والذي نرغب في اجرائه للتوصل الى المقاييس المختلفة التي نهدف اليها من وراء البحث .

نفترض اننا نريد دراسة وضع السكان بالنسبة للنشاط الاقتصادي ، فان احد اهدافنا من هـــذه الدراسة يكون معرفة عدد السكان الذين يكوّنوا القوة العاملة في الدولة متضمنة المشتغلين والمتعطلين وعــدد السكان الذين لا يعملون ولا يرغبون في العمل لاسباب مختلفة (أما لأنهم طلبة أو ربات منازل أو أصحاب أملاك يكتفون بايراداتهم من أملاكهم) وأخيراً عدد السكان الذين لا يستطيعون العمل (اما لأنهم أطفال صغار أو لأنهم مسنين فوق سن معين أو لانهم عجزة).اننا نوجه سؤالاً لكل فرد من السكان ونحصل منه على اجابة؛ الا ان هذه الاجابات المتفرقة في استمارات لا تساعدنا في الوصول الى هدفنا ، ولذلك يجب تصنيف هذه الاجابات في الفئات المختلفة التي نريدها ثم وضع نتائج التصنيف في جدول يظهر لنا الارقام المطلوبة ، هذه العملية نسميهـــا بالتصنيف والتبويب .

وهناك أنواع مختلفة للتبويب ولكنها جميعاً تهدف الى هدف واحد هو تجميع المعلومات في فئات وابرازها باكبر وضوح ممكن في أضيق حيز ممكن. وعهد اجراء هذه العملية لا يجب أن يكون غرضنا قاصراً على عرض البيانات عرضاً يتفق والبحث الذي نقوم به فقط ، خاصة اذا كان لهذه البيانات أوجه متعددة بحيث يمكن الاستفادة منها في أبحاث أخرى غير البحث الذي من أجله جمعت المعلومات (يحدث هذا غالباً في الابحاث الاحصائية الوصفية) . ولهذا يجب عرض البيانات عرضاً يساعد كل من يكون له غرض في البحث أن يجد ما يريده من معلومات تفيده . فالمعلومات الخاصة بالسكان في تعداد للسكان مثلا ، يجب تبويبها بحيث يمكن أن يستفيد منها الباحث الاقتصادي والاجتماعي والبيولوجي ... الخ . فالاقتصادي يستطيع أن يلجأ اليها لمعرفة عدد المشتغلين في أوجه النشاط الاقتصادي المختلفة ، والاجتماعي يجد فيها ما يساعده على دراسة التركيب العمري والجنسي للسكان بجانب توزيعهم تبعاً للمهن المختلفة ، والبيولوجي يستطيع أن يعتمد عليها في دراسة نمو السكان والتغيرات التي يمكن أن تصيب التركيب العمري من فترة الى اخرى .

ان الغرض من التصنيف هو كما ذكرنا تجميع الوحدات في فئات معينة وتختلف انواع هذه الفئات تبعاً لطبيعة البيانات المطلوب تبويبها وتبعاً للكيفية التي على أساسها سوف تستخدم بعد تبويبها . وأهم انواع التصنيف هي التصنيف الزمني وفيه تصنف المعلومات تبعاً لوحدات زمنية معينة كالشهر أو السنة مثلا ، والتصنيف الجغرافي وفيه تصنف المعلومات تبعاً لمناطق جغرافية معينة قد تكون هي الاقسام الادارية للدولة وقد تكون مناطق مختلفة في العالم، والتصنف الوصفي وفيه تصنف المعلومات تبعاً لاوصاف معينة تشترك فيها كل مجموعة من الوحدات مثلا تصنيف السكان تبعاً للمهن وتصنيف المؤسسات الصناعية تبعاً لنوع نشاطها ، والتصنيف الكمي وفيه تصنف المعلومات تبعاً لفئات رقمية معينة مثلا تصنيف الطلبة تبعاً لفئات أعمارهم حيث يمكن وضع الطلبة الذين في سن العشرين في مجموعة وهؤلاء الذين

في سن الواحدة والعشرين في مجموعة اخرى وهكذا . ويمكن توسيع مدى المجموعة فنقول عدد الطلبة الذين تتراوح أعمارهم بين ١٦ – ٢٠ سنة هم ١٠٠ طالب وعدد الذين تتراوح أعمارهم بين ٢٢ – ٢٤ سنة هم ٥٥ طالب وهكذا. ويسمى هذا التوزيع بالتوزيع التكراري وسوف نناقشه بتفصيل فيما بعد .

هذه هي الانواع المختلفة لتصنيف المعلومات من الناحية النظرية ، إلا أنه من الناحية العملية يمكن تصنيف الوحدات تبعاً لأساسين من هذه الاسس ، فنصنف السكان مثلاً تبعاً لاعمـــارهم في مناطق الدولة المختلفة او نصنف المؤسسات الصناعية تبعاً لنوع نشاطها وفئات حجمها أو نصنف الطلبة تبعاً لفئات أطوالهم وفئات أوزانهم فنقول مثلاً ان عدد الطلبة الذين تتراوح أوزانهم بين ٦٠–٦٥ كيلو والذين في نفس الوقت تتراوح أطوالهم بين ١٥٠–١٥٥سم هم ٣٠طالب، أو تصنف الاطفال تبعاً لفئات أعمارهم وفئات أوزانهم فنقول ان الاطفال الذين تتراوح أعمارهم بين ٢ – ٤ سنوات والذين تتراوح أوزانهم بين١٢–١٥ كيلو هم٥٠طفلا مثلاً. ومثل هذا النوع من التبويب يسمى بالتبويب المزدوج حيث يكون هناك أساسان لتصنيف الوحدات أي أن كل رقم في الجدول يدل على عدد من الوحدات ذات صفتين في نفس الوقف .

وليس هناك ما يمنع من تصنيف الوحدات تبعاً لأكثر من أساسين فنصف السكان مثلاً تبعاً لمكان سكنهم في الريف أو المدن وتبعاً لجنسهم ، ذكور أو اناث ، وتبعاً لحالتهم الزواجية أعزب أو متزوج أو مطلق أو أرمل ، أو نصنف المواليد تبعاً لجنسهم وتبعاً لفئات أعمار امهاتهم وتبعاً لشرعيتهم – مواليد شرعيين أو غير شرعيين . إلا أنه يجب أن نلاحظ انه كلما زادت عدد الاسس التي تصنف تبعاً لها الوحدات يصبح التبويب الناتج من عملية التصنيف معقداً يصعب فهمه ، ولذلك يحسن دائماً ألا نكثر في أسس التصنيف إلا إذا كانت الحاجة الماسة تدعو الى ذلك .

طرق التصنيف .

يمكن اجراء التصنيف بطرق مختلفة ، واختيار طريقة منها يتوقف على عوامل كثيرة منها عدد الوحدات التي نريد تصنيفها ومدى تنوع المعلومات الخاصة بهذه الوحدات. فكلما كان نطاق الدراسة ضيقاً كلما احتجنا الى اجراء التصنيف باليد حيث لا نحقق فائدة تذكر من استخدام الآلات وحيث يكلفنا استخدامها تكاليف كثيرة اذا قيست بالنسبة لنطاق البحث . أما اذا اتسع نطاق البحث بحيث أصبح يشمل وحدات كثيرة ومعلومات مختلفة يكون استخدام الآلات أمراً واجباً حيث نحقق فوائد كثير من السرعة في انجاز العمل وانقاص اليد العاملة التي نكون في حاجة اليها . وتظهر هذه الفوائد بشكل واضح في تعداد السكان حيث نحتاج الى معرفة النتائج في وقت يمكن فيه الاستفادة منها وحيث يصل عدد الوحدات الى الملايين .

التصنيف اليدوي :

لتوضيح طريقة التصنيف اليدي نفترض أن لدينا استمارات خاصة بخمسين مؤسسة صناعية ونريد تصنيفها تبعاً لموقعها . نبدأ باعداد الجدول الذي تجري عملية التصنيف فيه ونسميه جدول التفريغ وطبعاً يوضع في كعب الجدول مناطق الدولة المختلفة – مثلاً بيروت ، لبنان الشمال، لبنان الجنوب ، الجبل، البقاع . وبعد ذلك نتناول الاستمارات الواحدة تلو الاخرى ونقرأ البيان الخاص بموقع المؤسسة ، فاذا كان بيروت نضع أمامها اشارة معينة لتكن / . هذه الاشارة تدل على وجود وحدة (في هذه الحالة مؤسسة) في هذه الفئة للموقع. ونكرر هذه العملية حتى ننتهي من تفريغ البيان الخاص بموقع الخمسين مؤسسة . ولتسيل عملية عد هذه الاشارات يحسن أن نضع كل أربع اشارات بجوار بعضها هكذا ////،أما الاشارة الخامسة فتكتب بعكس اتجاه الاشارات الأربعة على هذا النحو ~~////~~. وبذلك يتكون لدينا مجموعات من الاشارات كل

منها يحتوي على خمسة مثلاً //متباعدة عن بعضها نوعاً ما وبذلك يمكن عدهـــا بسهولة (١٧ مؤسسة) . بعد ذلك يصمم جدول تنقل فيه المعلومات الموجودة في جدول التفريـغ ولكن بدلاً من الاشارات تكتب الارقام الدالة عليها وبذلك نكون قد أعددنا جدولاً يظهر توزيـع المؤسسات تبعاً لموقعها .

ويمكن أن يجري العمل بنفس الطريقة في حالة التبويب المزدوج والفرق الوحيد ان الاشارة توضع أمام صفتين تدلان على الوحدة التي نصنفها . ففي مثالنا السابق اذا كنا نصنف المؤسسات تبعاً لموقعهـا وفئات حجمها نبدأ باعداد جدول التفريـغ ونكتب في كعبه فئات الموقـع وفي عناوين أعمدته فئات الحجم التي نتفق عليها مثلاً ١ – ٤ ، ٥ – ٩ ، ١٠ – ١٩ ، ٢٠ – ٤٩ ، ٥٠ –٩٩ ، ١٠٠ أو أكثر . بعد ذلك نتناول الاستمارات الواحدة تلو الأخرى ونقرأ البيان الخاص بالموقع مثلاً بيروت ثم نقرأ البيان الخاص بعدد المشتغلين في المؤسسة مثلاً ٦٧ عامل فنضع الاشارة أمام بيروت وتحت فئة الحجم ٥٠ – ٩٩ وهكذا نستمر في وضـــع الاشارات بنفس الطريقة السابقة .

الا اننا قد نجد ان عدد الوحدات في كل استمارة من الاستمارات الخمسين كبيراً بحيث نضيع وقتاً طويلا اذا اردنا تمثيل **هذه** الوحدات بالاشارات السابقة وفي هذه الحالة يحسن كتابة العدد نفسه الخاص بعدد الوحدات ونضيف اليه العدد في الاستمارة التالية وهكذا . فــاذا أردنا تصنيف المؤسسات والمشتغلين فيها تبعاً لفئات الموقع فاننا نصمم جدول التفريـغ حيث نضع في كعبه فئات الموقع ثم نقسم جسم الجدول الى قسمين ، قسم خاص بتفريـغ المؤســات وقسم لتفريـغ المشتغلين. ثم نتناول الاستمارات الواحدة تلو الاخرى ونقرأ البيان الخاص بموقع المؤسسة فنضع الاشارة / أمام هذا الموقع ثم نقرأ البيان الخاص بعدد المشتغلين فتكتب العدد أمام نفس الموقع وهكذا .ويظهر ذلك في جدو التفريـغ الآتي : –

توزيـع المؤسسات والمشتغلين فيها تبعاً لموقع المؤسسة

الموقع	عدد المؤسسات	عدد المشتغلين
بـيروت	𝍸 //	١٠+٨+٥+١٨+٣٧+٢٣+٢
الشمال	///	٢٨ + ٣٥ + ١٥
الجنوب	////	٥ + ٢٢ + ٦٧ + ١٢
الجبـل	//	١٥ + ٩
البقاع	𝍸 /	٣٠+٧٢+٤٣+٢٥+١٩+٨

ويتضح من جدول التفريـغ السابق أن عـــدد الاشارات الخاصة بعدد المؤسسات يكون هو نفسه عدد الاعداد الخاصة بالمشتغلين وهذا أمر منطقي حيث أن كل اشارة تدل على مؤسسة وكل مؤسسة لا بد أن يكون فيها عدداً من المشتغلين . وواضح من الجدول اننا وجدنا ٧ مؤسسات في بيروت يشتغل فيها ٢ + ٢٣ + ٣٧ + ١٨ + ٨ + ٧ + ١٠ أي ١٠٥ مشتغل . وبذلك يكون متوسط عدد المشتغلين للمؤسسة الواحدة = $\frac{١٠٥}{٧}$ مشتغل وهكذا بالنسبة لمناطق الدولة الأخرى .

ويمكن اتباع نفس الطريقة إذا أردنا أن نصنف المؤسسات والمشتغلين تبعاً لموقع المؤسسة ونوع نشاطها أو تبعاً لموقع المؤسسة وفئة حجمها ، وفي هذه الحالة يتضمن كعب الجدول الأساسان اللذان نستخدمهما في التصنيف كما يظهر من الجدول التالي :

توزيـع المؤسسات والمشتغلين فيها تبعاً لموقع المؤسسة وفئة حجمها (مقـاساً بعدد المشتغلين)

الموقع وفئات الحجم	عدد المؤسسات	عدد المشتغلين
بيروت		
١ – ٤	/	٢
٥ – ٩	//	٥ + ٨
١٠ – ١٩	//	١٠ + ١٨
٢٠ – ٥٩	//	٢٣ + ٣٧
٥٠ – ٩٩		
١٠٠ وأكثر		
الشمال		
–		
–		
–		

التصنيف الآلي :

لاحظنا ان التصنيف اليدوي يؤدي الى ضياع الوقت والجهـد ولذلك لا يمكن الاعتماد عليه الا في البحوث التي يقل فيها عدد الوحدات وكمية البيانات

الخاصة بها . اما في البحوث الواسعة النطاق يتحتم علينا اجراء التصنيف باستخدام الآلات الاحصائية. واول ما يجب ان نلاحظه على الآلات الاحصائية انها لا تفهم البيانات الكلامية وانما تفهم الارقام فقط وفي شكل ثقوب . واستعمال هذه الآلات يقوم على الاسس الآتية :

١ – الاستعاضة عن الاستمارة ببطاقة خاصة من الورق المقوي يمكن ان تتداولها الآلات وتجري عملها عليها .

٢ – التعبير عن الوحدة بثقب معين امام رقم معين وفي عمود معين من البطاقة، فاذا كنا نريد تصنيف السكان تبعاً للجنس–ذكور واناث– وخصصنا العمود الثالث في البطاقة مثلاً لهذا البيان وعبرنا عن فئة الذكور برقم ١ وفئة الاناث برقم ٢ فاي ذكر من السكان سوف يعبر عنه بثقب في العمود الثالث امام رقم ١ في البطاقة . وواضح ان هذا الثقب يجري آلياً اي باستخدام آلة خاصة بهذه العملية .

٣ – فرز البطاقات بعد تثقيبها في آلة خاصة بذلك حيث تقوم هذه الآلة بتجميع البطاقات المثقوب فيها الرقم ١ في العمود الثالث مع بعضها وتجميع البطاقات المثقوب فيها الرقم ٢ في العمود الثالث مع بعضها :

٤ – توضع البطاقات بعد فرزها في آلة اخرى تقوم بعد البطاقات في كل مجموعة وتسجل لنا العدد امام رقم ١ اي الذكور وامام الرقم٢ اي الاناث.

٥ – واخيراً تؤخذ هذه النتائج المسجلة وتوضع في شكل جدول .

وبذلك عندما نقرر استخدام الآلات الاحصائية يتعين علينا اجراء عمليات معينة هي اعداد الدليل ، والترميز ، ومراجعة الترميز ، والتثقيب ، ومراجعة التثقيب ، والفرز والتبويب ، كما يتعين علينا استخدام الآلات التي تساعدنا في هذه الاعمال المختلفة (ابتداء من عملية التثقيب) .

اعداد الدليل :

يجري الترميز للبيانات بناء على دليل موضوع ومعـــد لهذا الغرض . والمقصود بالدليل القائمة التي تحتوي على الفئات المختلفة الخاصة ببيان ما والأرقام التي ترمز إلى هذه الفئات فنقول مثلا الدليل الخاص بالمهن أي القائمة التي تحتوي على المهن المختلفة التي يمكن أن توجـد في مجتمع ما والارقام الرمزية التي تدل على هـذه المهن عند اجراء العمليات الخاصة بالتبويب . كذلك نقول الدليل الخاص بالنشاط الاقتصادي أي القائمة التي تشتمل على مختلف أنواع النشاطات الاقتصادية والأرقام الرمزية التي تدل عليها . كذلك نقول الدليل الخـــاص بالتقسيم الاداري الدولة أي القائمة التي تشتمل على التقسيمات الادارية المختلفة في الدولة والارقام الرمزية الدالة عليها ... الخ .

ولاعداد دليل خاص ببيان ما يجب أن تتبع الخطوات التالية :

١ – يجب أن يكون واضحاً لدينا المعنى المقصود من البيان والغرض الذي نتوخاه احصائياً من السؤال عن هذا البيان . فاذا أردنا مثلا أن نعد دليلاً للتقسيم الاداري للدولة يجب أن نفهم جيداً المقصود بهذا التقسيم وهل نحتاج في دراستنا إلى تقسيم اداري مفصل أو تقسيم عام قائم على أساس الأقسام الرئيسية في الدولة ، ويتضح من هذا المثال العلاقة بين الترميز والتبويب حيث أن البت في هذه المشكلة يتوقف على الجداول الاحصائية المعدة لتصنيف البيانات فيها. كذلك يتضح العلاقة بين الترميز وتصميم الاستمارة والتعليمات إلى العدادين حيث إذا كان المطلوب تصنيف البيانات تبعاً للأقسام العامة يمكن ان يكتفي بتسجيل عنوان وحدة الاستبيان على أساس القسم الاداري الرئيسي ، أما اذا كان المطلوب التصنيف تبعاً للأقسام الفرعية يتحتم تسجيل العنوان مفصلاً . وإذا أردنا مثلاً أن نعد دليلاً للمهن يجب أن نفهم أولاً المقصود بالمهنة وهل نحتاج في دراستنا إلى تصنيف عـــام للمهن أو تصنيف للمهن على أساس عام ومفصل في نفس الوقت . وبالطبع يترتب على هذا البحث أن تؤخذ العلاقة بين التبويب والترميز والاستمارة وتعليمات العدادين في الاعتبار .

٢ – بعد أن نخلص من الخطوة الأولى تكون الخطوة الثانية هي حصر جميع الحالات التي يمكن أن نحصل عليها بتوجيه السؤال المعد في الاستمارة إلى جميع المستجوبين موضوع الدراسة . ويمكن أن يكون هذا أمراً سهلاً كما هو الحال بالنسبة للأسئلة الخاصة بالبيانات البسيطة مثل الجنس والحالة الزواجية والحالة التعليمية و .. الخ ، ويمكن أن يكون هذا الحصر أمراً صعباً كما هو الحال بالنسبة للمهن والنشاطات الاقتصادية والأمراض التي يمكن أن يصاب بها أفراد المجتمع موضوع الدراسة ونظراً لصعوبة اعداد الدليل الخاص بهذه البيانات قامت المنظمات الدولية المعنية بشؤون الاحصاء بوضع دليل لكل من هذه البيانات يمكن أن يعتمد عليه .

٣ والخطوة الثانية هي وضع أرقام تدل على الحالات التي تم حصرها في الخطوة السابقة

مثال :

في دراسة عن الصناعة اشتملت الاستمارة على سؤال عن الكيان القانوني للمؤسسات الصناعية والمطلوب اعداد دليل خاص بهذا البيان . ان الخطوة الأول كما قدمنا هي فهم المقصود بالكيان القانوني ثم حصر جميع الحالات الخاصة بهذا الكيان وهي :

مؤسسات فردية
مؤسسات تضامن
مؤسسات توصية بسيطة
مؤسسات توصية بالاسهم
شركات مساهمة الخ .

بعد ذلك يعطي كل من هذه الكيانات القانونية رقم خاص فيتكون لدينا قائمة بالدليل المطلوب يكون كالآتي :

مؤسسات	فردية	١
مؤسسات	تضامن	٢
مؤسسات	توصية بسيطة	٣
مؤسسات	توصية بالأسهم	٤
شركات	مساهمة	٥

يعطي لنا هذا المثال فكرة بسيطة عن اعداد الدليل ، إلا اننا في هذه العملية يمكن أن نواجه مشاكل مختلفة تجعل العمل أكثر صعوبة . والمشاكل التي نواجهها في هذا الصدد يمكن تلخيصها في الآتي :

أولاً : صعوبة حصر جميع الحالات الخاصة ببيان ما ، وقد سبق أن أشرنا إلى هذه المشكلة فيما يختص بالمهن والنشاط الاقتصادي والأمراض وأسباب الوفاة والبضائع التي تدخل في التجارة الخارجية . لذلك أعدت اللجنة الاحصائية للمجلس الاقتصادي والاجتماعي بالامم المتحدة دليلا للنشاط الاقتصادي وبضائع التجارة الخارجية – كما أعد مكتب العمل الدولي دليلاً للمهن ، وكذلك أعدت منظمة الصحة العالمية دليلاً للامراض والايذاءات واسباب الوفاة . وكل دليل دولي يشتمل على حصر شامل للحالات التي يمكن أن توجد فيما يتعلق بالبيانات الخاصة بهذا الدليل ، وكذلك على ترميز لهذه الحالات . كذلك أعد كل دليل منها بحيث يمكن أن يستخدم على أساس أن المطلوب اجراء تصنيف تفصيلي يشتمل على الأقسام الرئيسية وما يتفرع عن هذه الأقسام من حالات فرعية وبذلك لا تضيع العلاقة بين هذين النوعين من التصنيف . كما ان كل دليل منها قد أعد بحيث يكون مرنا أي يمكن أن يتسع لادخال حالات جديدة إذا ظهرت مثل هذه الحالات في المستقبل وبذلك لا يستدعي الأمر تعديل الدليل من وقت إلى آخر .

وبالنسبة لهذه المشكلة يمكن أن أورد الملاحظات التالية :

أ – في بعض الأحيان لا يكون تفصيل الحالات الخاصة ببيان ما أمر

مرغوب فيه إذا اخذ في الاعتبـــار التبويب النهائي وبذلك يمكن اعطاء الحالات المطلوب ظهورها في التبويب ارقاماً رمزية ثم تجمع بقية الحالات في فئة واحدة دون تفصيل وتعطي جميعاً رقماً رمزيـــاً واحداً . ففي مثالنا السابق إذا كان الاهتمام موجه إلى معرفة عدد المؤسسات الفردية والتضامن والمساهمة فقط فيكون الدليل كالآتي :

مؤسسات فــردية	١
مؤسسات تضامن	٢
مؤسسات مساهمة	٣
مؤسسات أخرى	٤

وفئة المؤسسات الأخرى يقصد بها كل أنواع الكيان القانوني التي لم يخصص لها بند مستقل في الدليل . ونظراً لأنه من الصعب في بعض الأحيان تحديد جميع الحالات التي يمكن أن توجد فان فئة اخرى مثل مؤسسات أخرى أو غراض أخرى تكون ذات فائدة قصوى حيث انها فئة مطلقة يمكن أن تستوعب كل الحالات التي لم يخصص لها بند مستقل أو التي لم يتوقع وجودها مصمم الدليل ، وبذلك يكون الدليل المنشأ مغطيـــاً لجميع الحالات وهو أساس يجب أن ننظر اليه كحجر الزاوية عند انشاء اي دليل .

ب – وحيث ان كثيراً ما ترد الاستمارات إلى المكتب دون أن يستوفي بعض بياناتها لأي سبب من الأسباب ، ونظراً لصعوبة الرجوع إلى الميدان لاستيفائها – واظهار هذه الحالات في التبويب النهائي أمر بالغ الأهمية لأنه يمكن أن يكون مقياساً لمدى وعي المستجوبين ولمدى الدقة في جمع المعلومات – لذلك يجب أن يخصص في الدليل رقما ومزياً للدلالة على كل بيان لم ترد الاجابة عنه ، وعادة يخصص آخر رقم في حدود الخانة لذلك الدليل مثلاً :

ذكر ١

انثى ٢

لم يبين ٩

وقد خصص الرقم ٩ لفئة لم يبين لأنه آخر رقم في خانة الآحاد وهي الخانة التي يتكون منهـــا هذا الدليل ، واذا كانت أرقام الدليل تتكون من خانتين فيخصص لفئة لم يبين الرقم ٩٩ لأنه آخر رقم في حدود خـــانة العشرات .

ج - يمكن أن يكون التفصيل في الدليل أمر مرغوب فيه ولكن يجب أن يؤخذ في الاعتبار عند انشاء دليل مفصل يحتوي على عـــدد كبير من الحالات عدد الخانات التي سوف يشغلها الدليل بناء على هذا التفصيل نظراً لأن لعدد الخانات التي يشغلها الدليل أهمية كبيرة من الناحية العملية التطبيقية وكذلك من ناحية بطاقة الآلات التي سوف ينقل اليها الدليل بالتثقيب . وقد تكون هناك رغبة في تقليل عدد الحالات الخاصة بدليل ما تجنباً للتعقيد في العمل عند مطالبة المرمزين بتطبيقه وتوفيراً للوقت والجهد والنفقات بتقليل عدد الخانات في بطاقة التثقيب ومـــا يلي ذلك من خطوات حتى اخراج التبويبات بواسطة الآلات ... الخ . لهذا يجب أن نقارن دائماً بين الفائدة التي تتحقق من التفصيل في الدليل وبين الجهد والوقت والنفقات وتعقيد العمل الذي يتطلبه تطبيق دليل مفصل . ولتجنب التفصيل في دليل ما ، تعطى الحالات الاساسية أرقاماً رمزية وتجمع الحالات الاخرى في فئة واحدة تسمى بحالات اخرى مثل حالة (مؤسسات اخرى) التي سبق الاشارة اليها .

د - كذلك لا بد عند وضع الدليل أن نراعي توفر المرونة فيه - وقد سبق أن أشرت الى ذلك فيما يختص بالتصنيفات الدولية . ويقصد بالمرونة أن يكون في الدليل موضع لاضافات جديدة دون أن تحدث تلك الاضافات قلقلة في مبنى الدليل . على انه يجب أن نلاحظ أن البيانات التي يوضع لهـا

دليل تختلف فيما بينها من ناحية حاجتها إلى المرونة فهناك بيانات بطبيعتها لا تستدعي ذلك حيث لا مجال فيها لأي حدث ، ومثال ذلك البيان الخاص بجنس السكان (ذكر أو انثى) . وهناك بيانات بطبيعتها عرضة لهذا المستحدث مثل البيان الخاص بالمهنة أو النشاط الاقتصادي أو التوزيع الجغرافي ... الخ .

مثـــال :

إذا أردنا وضع دليل للتقسيمات الادارية في مصر نلاحظ أن هناك محافظات ومديريات ويجب أن يوضح الدليل إذا كانت هذه المديريات في الوجه البحري أو في الوجه القبلي . فعند وضع الدليل لهذه التقسيمات الادارية يجب أن يتجه اهتمامنا الى المستقبل حيث يمكن أن تنشأ محافظة جديدة أو مديرية جديدة سواء في الوجه القبلي أو البحري ولذلك يجب أن يسمح الدليل الموضوع بادراجها فيه دون أن يهدم مبناه . وتطبيقاً لذلك يمكن أن يخصص العشرة أرقام الأولى للمحافظات وان كان عددها ٤ فقط ، ويخصص العشرين رقمـــاً التالية لمديريات الوجه البحري وان كان عددها ٨ وبذلك يكون الدليل كالآتي :

اول محافظة	رقم ١	وآخر محـافظة	رقم ٤
اول مديرية في الوجه البحري	رقم ١١	وآخر مديرية	رقم ١٨
اول مديرية في الوجه القبلي	رقم ٣١	وآخر مديرية	رقم ٣٨
اول أقسام الحدود	رقم ٥١	وآخر قسم	رقم ٥٥

فإذا أنشأت محافظة جديدة تعطى رقم ٥ وإذا أنشأت مديرية جديدة في الوجه البحري تعطى رقم ١١ وهكذا . وبالطبع نلاحظ أن المدى (عشرة أو عشرين أو أكثر) يتوقف في مثل هذه الحالة على عدد المفردات التي يتوقع زيادتها .

مثـــال ٢ :

في التصنيف الخاص بالمهنة أو بالنشاط الاقتصادي او بالتجارة الخارجية يجري التقسيم على أساس أقسام وكل قسم يقسم الى ١٠ مجموعات رئيسية مع أن بعض الاقسام يمكن ان لا تحتوي إلا على مجموعة واحدة أو مجموعتين فقط في الوقت الحاضر أما باقي الارقام العشرة فتكون كفراغ يمكن أن توضع فيه أي مجموعة جديدة في المستقبل . وكذلك كل مجموعة رئيسية يمكن أن تقسم إلى عشرة مجموعات فرعية مع انها في الحاضر قد لا تحتوي إلا على عدد أقل من عشرة مجموعات فرعية ويكون باقي الأرقام كفراغ توضع فيه أية مجموعات فرعية تنشأ في المستقبل .

وفي هذا الصدد نناقش مشكلة انشاء دليل لأقسام وفروع متفرعة عنها وتابعة لها . في هذه الحالة يخصص الرقم الأول للدلالة على القسم والرقم الثاني للدلالة على الفرع وإذا كان هناك ما يتفرع عن هذه الفروع يخصص الرقم الثالث للدلالة عليها . وبالنسبة لجميع الفروع المتفرعة عن قسم ما يكون دليلها جميعاً مشتركاً في الرقم الأول الدال على هذا القسم – مثلاً القسم رقم صفر وما يتفرع عنه يكون الدليل الخاص به ٠٠ ، ٠١ ، ٠٢ ، ٠٣ وهكذا ، والقسم رقم ١ وما يتفرع عنه يكون الدليل الخاص به ١٠ ، ١١ ، ١٢ ، ١٣ وهكذا ، والقسم رقم ٢ وما يتفرع عنه يكون الدليل الخاص به ٢٠ ، ٢١ ، ٢٢ ، ٢٣ ؛ فإذا كان لدينا رقم دليل ٢٣ فيقرأ كالآتي : الفرع رقم ٣ التابع للقسم رقم ٢ ، وإذا كان لدينا رقم دليل ٤٧ فنقرأه الفرع رقم ٧ التابع للقسم رقم ٤ ، وإذا كان لدينا رقم دليل ٥٨٢ فنقرأه الفرع رقم ٢ التابع للفرع رقم ٨ التابع للقسم رقم ٥ (يمكن الاطلاع على التصنيفات الدولية لأخذ فكرة تفصيلية عن هذا الموضوع) .

ولا شك أن ترك الفراغ في الدليل حتى يمكن اضافة بنود جديدة إذا استحدثت مثل هذه البنود أمر يحقق التناسق في الدليل . ففي المثال الخاص

بالتقسيمات الادارية كان من الممكن أن تعطى المحافظات الارقام من ١ الى ٤ ومديريات الوجه البحري تبدأ بالرقم ٥ وتنتهي بالرقم ١٢ ومديريات الوجه القبلي تبدأ بالرقم ١٣ وتنتهي بالرقم ٢٠ وأقسام الحـــدود يبدأ الرقم ٢١ وتنتهي بالرقم ٢٥.وإذا استحدثت تقسيمات ادارية جديدة تعطي الرقم التالي ٢٦ سواء كانت تلك الاضافة محافظة أو مديرية او قسم حدود ، ولكننا بذلك نجعل الدليل غير متناسق وغير منطقي ، الامر الذي يجعل فهمه وتطبيقه أمراً صعباً . على أن ذلك لا يحدث إلا بالنسبة للبيانات المقسمة إلى أقسام تختلف فيا بينها ، إلا انه هناك بيانات ليست في حاجة الى مثل هذه المعالجة لأنه لا يترتب أي ضرر أن يدرج البند المستحدث في نهاية القائمة المعدة قبل استحداث هذا البند مثل البيان الخاص باسباب التعطل .

ثانياً : لا تنتهي عملية اعداد الدليل بحصر الحالات واعطائها أرقام تدل عليها ، إذ يجب توزيع البيانات على بطاقة التثقيب . والقاعدة الأساسية في هذا الصدد هي تخصيص عدد من الأعمدة في البطاقة لبيان ما تبعا لعدد الخانات التي يتكون منها الدليل الخاص بهذا البيان . وبذلك يتحدد مجموع الأعمدة التي يحتاجها تثقيب الأرقـام الرمزية الخاصة بعدة بيانات جمعت في دراسة احصائية ما (تعداد السكان مثلاً) وبالتالي يتحدد عدد بطاقات التثقيب التي نحتاجها ويتوقف ذلك طبعاً على نوع البطاقة المستخدمة ، ويتحدد هذا النوع تبعاً لنوع الآلات الاحصائية الموظفة في العمل فإذا كان مجموع الأعمدة التي نحتاجها ٨٠ أو أقل وكان عدد الأعمدة في البطاقة ٨٠ فاننا نحتاج بطاقة واحدة لكل وحدة ، أما إذا كان عدد الأعمدة التي نحتاجها أكثر من ٨٠ فاننا بذلك نحتاج لبطاقتين لكل وحدة وهكذا .

والمشكلة الثانية في هذا الصدد هي توزيع البيانات على بطاقة التثقيب ، بيان مثلاً يحتاج لعمودين هل يكون هذين العمودين في أول البطاقة أو في وسطها أو في آخرها . ويحتاج البت في هذه المشكلة الى خبرة فنية بالآلات الاحصائية المستخدمة ولا مجال هنا لبحث هذا الموضوع .

الترميز :

يقصد بعملية الترميز كتابة الرموز الرقمية أمام اجابات الأسئلة المختلفة طبقاً للدليل الذي أعددناه ، وبذلك اذا كانت بعض المعلومات في الاستمارة عبارة عن أرقام (عدد العمال مثلاً) فلا يكون هناك حاجة لترميزها وانما يحسن احاطة الرقم بدائرة بلون خاص حتى يتنبه المثقب اليها .

وهناك طرق مختلفة للترميز ، حيث يمكن أن تصمم اسئلة الاستمارة بحيث يكتب أمام كل سؤال الاجابات المحتملة له وكل اجابة لها رقم خاص بهـا يكون كرمز لها ، وبذلك عندما يؤشر المستجوب او العداد أمـام احدى هذه الاجابات يكون في نفس الوقت قد حدد الدليل الرمزي الخاص بالاجابة ، وبهذه الطريقة نختصر اجراء عملية الترميز . بالرغم من الميزة الكبيرة لهـذه الطريقة لا تتبع الا نادواً حيث أن اتباعها يحتاج الى تحديد جميع الاجابات المحتملة لكل سؤال تحديداً دقيقاً حتى لا نواجه اجابة ما ليس لها مكان في الاستمارة، الأمر الذي يضطرنا في هذه الحالة الى اضافة فئة بعنوان « آخرين» الا أن هذه الفئة قد تتضخم بوقوع عدد كبير من الاجابات فيها وبذلك لا يساعد تبويب هذا البيان على اجراء التحليل الاحصائي الدقيق . وهناك عيب آخر لهذه الطريقة حيث ان اتباعها يستدعى استخدام فئات عامة وبذلك تضيع تفاصيل اجابة المستجوب ، الأمر الذي لا يساعد في اكتشاف أي خطأ وقع فيه العداد . واذا كان العداد مطالباً باجراء بعض العمليات الحسابية قبل تحديد اجابة المستجوب ضمن الفئات المطبوعة في الاستمارة فان أي خطأ يقع فيه العداد لا يمكن اكتشافه فيما بعد . هذا فضلاً عن أن خطأ العداد في وضع اجابة المستوجب أمام الفئة الخاصة به لا يمكن اكتشافه عندما ترد الاستمارات الى المكتب .

والطريقة الثانية للترميز هي ترميز البيانات على الاستمارة نفسها بعد جمع المعلومات ومراجعتها. واستخدام هذه الطريقة يساعدنا في توفير الوقت اللازم

لنقل البيانات بعد ترميزها الى كشف خاص بالترميز كما انه يمكن تفادي الأخطاء التي تحدث عند النقل . والعيب الأول لهذه الطريقة هو عدم توفر الفراغ اللازم لكتابة أرقام الدليل حيث أن الاستمارة تسجل فيها المعلومات وملاحظات العدادين ثم يجري عليها التصحيح في الميدان ثم في المكتب، الأمر الذي يجعلها تكتظ بالكتابة . على أن هذه الصعوبة يمكن التغلب عليها بتخصيص فراغ خاص للترميز في الاستمارة نفسها وتنبيه العدادين الى عدم كتابة أي شيء في هذا الفراغ وكذلك يمكن أن يستخدم حبر بارز اللون عند اجراء الترميز . والعيب الاساسي لهذه الطريقة هو أن ترتيب البيانات في الاستمارة قد لا يتفق مع ترتيب البيانات في بطاقات التثقيب حيث يكون للآلات أثر كبير في توزيع البيانات على أعمدة هذه البطاقات ، هذا فضلاً عن كثرة الاستمارات المرمزة التي تذهب الى المثقبين وتكون سبباً في ارتباكهم في العمل .

والطريقة الثالثة هي ترميز البيانات في كشوفات تصمم خصيصاً لهذا الغرض ، ويمكن تصميم هذه الكشوفات بحيث ننقل البيانات اليها في شكل رموز رقمية تكتب في اعمدة معنونة بعناوين البيانات المختلفة ويمكن ترتيب هذه الاعمدة بحيث تتفق مع الترتيب اللازم لاجراء التثقيب بسرعة وهو الترتيب الذي يتفق تماماً مع ترتيب الاعمدة في بطاقة التثقيب حسب تصميمنا لها .

والطريقة الرابعة هي ترميز البيانات في بطاقة ترميز خاصة بكل حالة على حدة ، وبذلك تختلف عن الطريقة السابقة حيث انه في هذه الطريقة يكون لكل وحدة بطاقة ترميز خاصة بها بينما في الطريقة السابقة ترمز البيانات الخاصة بعدة وحدات في كشف واحد ويتوقف عدد الوحدات على اتساع الكشف الذي نستعمله . والميزة الاساسية لهذه البطاقة انه يمكن استخدامها كبطاقة أرشيف نرجع اليها كلما اردنا تصنيف خاص للبيانات لم يجرى من قبل.

والعيب الاساسي لهذه الطريقة هو ما نتحمله من تكاليف عندما نقرر استخدامها حيث نحتاج لعدد كبير منها خاصة في الدراسات الواسعة النطاق.

على أنه يمكن الجمع بين بعض هذه الطرق حيث تصمم الاسئلة ذات الاجابات المحددة بحيث تشتمل على الترميز الخاص بها على الاستمارة نفسها بينما تترك الاسئلة الاخرى دون ترميز على الاستمارة على أن يجري ترميز اجاباتها في المكتب بعد ذلك اما على الاستمارة نفسها أو في كشف ترميز مفصل أو في بطاقات ترميز .

ملاحظات على إعداد الدليل والترميز :

١ – يحسن تصميم الدليل الخاص ببيان ما على أساس تصنيف حالات هذا البيان أكثر تفصيلاً من التصنيف الذي سوف يظهر في التبويب النهائي ،حيث لا يمكن الحكم على مدى اتساع الفئات او ضيقها الا بعد ظهور الجداول الأولى . فإذا تبين لنا من هذه الجداول المعدّة على أساس التصنيف المفصل ان الفئات المستخدمة واسعة دون مبرر نظراً لوجود عدد قليل من الحالات في كل منها فإنه من السهل تجميع الفئات لإظهارها في الجداول النهائية.والعكس إذا تبين ازدحام الفئات بالتكرار فإنه يكون من الصعب تقسيم الفئات ويتحتم علينا في هذه الحالة إعادة العمل مرة اخرى على أساس أكثر تفصيلاً . لهذا وحتى نضمن لأنفسنا خط الرجعة يحسن التصنيف على أساس تفصيلي إلى أقصى درجة ممكنة .

٢ – إذا كان التبويب سوف يجري يدوياً يحسن الابتعاد عن الرموز الرقمية حيث من السهل الخطأ فيها . وفي هذه الحالة يمكن أن تكون الرموز عبارة عن حروف أبجدية لها علاقة بالكلمات التي تمثلها أو اختصارات لهذه الكلمات واستخدام هذه الطريقة يجنبنا كثيراً من الخطأ . وإذا أردنا استخدام الرموز كدليل يحسن تجنب الرموز التي يمكن ان تختلط فيما بينها عند كتابتها.

٣ — اذا كانت بعض البيانات عبارة عن أرقام (عدد العمال مثلا) لا نحتاج الى توقيع دليل لها ويحسن إحاطة الرقم بدائرة بلون خاص حتى يتنبه المثقب اليها .

٤ — يجب مراجعة الترميزقبل البدء في العملياتالآلية الأخرى (التثقيب والفرز والتبويب) .

٥ — إذا كان البيان يحتاج لإجراء عمليات حسابية قبـل توقيع الدليل الخاص به يحسن الفصل بين العمليتين فيقوم موظف باجراء العمل الحسابي ومراجعته أولاً ثم يقوم موظف آخر بتوقيع الدليل .

٦ — يحسن أن يكون المرمزين ملمين بالتعاريف والاصطلاحات المستخدمة في الدراسة ولذا يكون من الافضل توظيفهم من العدادين أو المراجعين الذين سبق لهم العمل في هذه الدراسة او من موظفي المكتب الذين لهم دراية بهذه التعاريف والمصطلحات .

٧ — يجب أن ينظم العمل في الترميز بحيث يجري العمل روتينياً ، فإذا وجدت مشاكل خاصة بالعمل يحسن ان يكون هناك موظف مختص لمعالجتها.

٨ — إذا احتوت استمارة البحث على بيانات صعبة الترميز (ترميز المهن أو النشاط الاقتصادي مثلا) يكون من الافضل تقسيم عملية الترميز بحيث تعطى البيانات الصعبة لموظفين متخصصين في ترميزها وتعطى البياناتالأخرى لمرمزين آخرين ، وفي هـذه الحالة يحسن ان يقوم بمراجعة الترميز ككل موظف واحد .

٩ — يجب أن تتم عملية الترميز بناء على تعليمات موجهة للمرمزين حتى يمكن توحيد خطوات العمل وطريقته بين جميع المرمزين القائمين بالعمل .

١٠ — ان دقة الترميز تتوقف أساساً على الدقة التي جمعت بها البيانات في استمارة البحث وعلى التوضيح الذي تسجل به تلك البيانات ، خصوصاً بالنسبة لبيانات

كتلك الخاصة بالمهنة او بالنشاط الاقتصادي. فإذا كنا بصدد النشاط الاقتصادي مثلاً يجب ان يكون البيان المسجل في الاستمارة موضحاً الى أبعد درجة حتى لا يختلط الأمر على المرمزين ، فلو وردت في الاستمارة عبارة « ملابس » لما كان ذلك كافياً لتوقيع الدليل الصحيح اذ يحتمل ان تكون المؤسسة مصنعاً للملابس أو تجارة ملابس بالجملة أو تجارة ملابس بالمفرق . وكل من تلك الاحتمالات يرشحها لدليل مختلف كل الاختلاف عن الآخر .

١١ – يجب أن يكون واضحاً للقائمين بالعمل في الترميز مدى الخطورة التي تترتب على عدم الدقة في ترميز البيانات حيث يكون ذلك بمثابة تغيير للبيانات ، الامر الذي يجعل النتائج غير دقيقة ومضللة . وواضح ان معنى ذلك القضاء على كل المجهودات التي سوف تبذل عقب توقيع الدليل حتى ظهور النتائج .

اجراءات تنظيمية لعمليات الترميز :

يجب ان تكون الاجراءات اللازمة لتنظيم عملية الترميز قد أعدت وجهزت قبل البدء في العمل ، والاجراءات اللازمة هي :

١ – إعداد المطبوعات اللازمة للعملية – يجب إعداد نسخاً من الدليل الذي سوف يستخدم في العملية ، وتوزيعه على المرمزين ، ويحسن ان يشتمل الدليل المطبوع على مقدمة تشرح بنود الدليل المختلفة واذا كانت التصنيفات الدولية سوف تستخدم في الترميز (للمهن او النشاط الاقتصادي .. الخ) فيجب إعداد نسخاً منها مع مقدمة تشرح كيفية استعمالها في الترميز ، ويحسن أن تحتوي هذه النسخ على فهرس حسب الحروف الهجائية لمساعدة المرمزين في العمل نظراً لتعقيد هذه التصنيفات وطولها . وأثناء العمل إذا أجريت بعض التعديلات أو إذا ظهرت بعض المشاكل التي أثارها المرمزين وقدمت لها الحلول يحسن إعداد مطبوعات بذلك وتوزيعها على الموظفين أولاً بأول . ويا حبذا

لو كان هناك تنظيم يمكن بمقتضاه أن يحتفظ كل مرمز بالمطبوعات الخاصة بالعمل في ملف خاص به .

٢ – اختيـــار الموظفين لعملية توقيع الدليل – يجب اختيار الموظفين الصالحين لهذه العملية وقد سبق أن أشرت الى انه يحسن ان يكـــونوا ملمين بالتعاريف والمصطلحات وكل الاسس التي يجري عليها العمل . وبالاضافة الى ذلك يجب ان يتوفر في الموظفين الذين يتم اختيارهم للترميز بعض الصفات مثل قوة الذاكرة وعدم النسيان والقدرة على التصرف وخاصة بالنسبة للموظفين الذين سوف يقوموا بالعمل مستخدمين أنواع الدليـــل المعقدة ، وكذلك يجب ان يتحلوا بالدقة وعدم الاستهتار . وتجري المفاضلة دائما بـــين اي الجنسين – الذكور او الاناث – اصلح لهذه العملية ويفضل البعض الاناث لتحليهن بالصبر وقدرتهن على الاعمال الرتيبة المكررة ، الا ان الذكور اثبتوا كذلك صلاحية في هذا العمل وخاصة بالنسبة للتصنيفات المعقدة التي تحتاج الى نوع من المرونة . لذلك عند بحث هـــذا الموضوع يجب أن لا يكون على أساس مقدرة كل من الجنسين وإنمـــا على أساس ظروف كل بيئة وتقاليدها وتوفر العدد اللازم من الجنسين وحاجة كل منهما للعمل .

٣ – تدريب الموظفين على توقيع الدليل – يجب تدريب الموظفين على العمل الذي سوف يوكل اليهم ، ويبدأ التدريب نظريا باعطاء الدروس لشرح الدليل وطريقة توقيعه وكيفية حفظه بسهولة . ومدة هذا التدريب تختلف تبعاً لنوع الدليل المطلوب التدريب عليه من ناحية التعقيد او البساطة . وبعد التدريب النظري يبدأ التدريب العملي بتوزيع استمارات بهـــا البيانات التي سوف ترمز على الموظفين ومطالبتهم بترميزها تحت اشراف المدربين . وبالطبع لا يجب استخدام الاستمارات الاصلية حيث ان الموظفـــين سوف يرتكبون أخطاء كثيرة عندما يبدأون في العمـــل ولذلك يجب أعداد كشوف لهذا

الغرض ويكفي أن تحتوي على البيانات موضوع الترميز فقط . وعلى ضوء نتائج التدريب يمكن استبعاد من لا يصلحون لهذا العمل .

التعليمات الخاصة بالترميز :

سبق أن أشرت الى أن عملية الترميز يجب أن تجري بناء على تعليمات مفصلة وواضحة لتوحيد خطوات للعمل بين الموظفين. ويحسن أن تشتمل هذه التعليمات على الآتي : —

١ — شرح الغرض من الترميز وموقع عملية الترميز من العملية الاحصائية بكاملهـــا .

٢ — بيان بضرورة اطلاع المرمزين على التعليمات الموجهـــة للعدادين والمراجعين والعاملين على الآلات الثاقبة والفارزة (ولهذا يحسن توزيـع نسخاً من هذه التعليمات على المرمزين كذلك) .

٣ — الترميز يكون للبيانات بعد تصحيحها سواء أجرى التصحيح من قبل العداد أو المشرف أو المراجع . أي أن لا يجب الالتفات الى البيانات الأصلية الخاطئة .

٤ — تحديد بيانات الاستمارة التي سيوقع لها الدليل .

٥ — تحديد أنواع الدليل التي سوف تستخدم في العمل (مـع ملاحظة ضرورة توفر نسخة من كل دليل يجري استخدامه مع كل موظف) .

٦ — شرح لكيفية تسجيل ارقام الدليل ، وهذا يشتمل على بيان المكان الذي يوقع فيه الدليل هل على يسـار أم على يمـين البيان ، أم فوقه أم تحته أم في خانة مخصصة لذلك . كذلك يجب تحديد لون القلم الذي يستخدم في توقيـع الدليل (وعادة يستخدم اللون الازرق) ولون القلم الذي يستخدم في مراجعه الترميز (وعادة يستخدم اللون الاحمر) . كذلك يجب التنبيه على

الكتابة بخط واضح مقروء واذا حدث خطأ يشطب الرقم الخاطىء ويكتب فوقه الرقم المصحح ، أي لا يجب مسح رقم الدليل او تغييره عند حدوث خطأ . كذلك يجب الاشارة الى البيانات التي لن يجري ترميزها وانما يجب احاطتها بدائرة لتنبيه المثقبين اليها .

٧ – تحديد ترتيب العمل بالنسبة لترميز البيانات المختلفة التي تحتوي عليها الاستمارة . ويدور نقاش حول الطريقة التي يقوم بها المرظف بتوقيع الدليل، هل يقوم بترميز بيان ما بالنسبة لجميع الوحدات ثم ينتقل الى بيان آخر أم يقوم بترميز جميع البيانات الخاصة بوحدة ما ثم ينتقل الى ترميز بيانات وحدة أخرى .

ان كلا من هاتين الطريقتين لها مزاياها وعيوبها . ففي الحالة الاولى يقل المجهود الذي يبذله الموظف حيث انه يركز ذهنه في نوع منواحد من البيانات فيتعود عليه ويصبح من السهل توقيع رقم الدليل دون الحاجة الى البحث عنه طويلاً . الا أن الموظف لا يستطيع الربط بين بيانات الوحدة الواحدة ، لذلك لا تساعد هذه الطريقة على اكتشاف الأخطاء التي يمكن أن توجد في البيانات بالرغم من مراجعتها قبل ذلك .

أما بتطبيق الطريقة الثانية يمكن الربط بين بيانات الوحدة الواحدة ويستطيع الموظف تبعاً لذلك الاهتداء الى الدليل الصحيح لبيان ما قد لا يمكن الاهتداء اليه دوو هذا الربط . كذلك فان الربط بين البيانات يساعد في اكتشاف الأخطاء التي يحتمل وجودها فيها . إلا انه في هذه الحالة يبذل الموظف مجهوداً كبيراً يتمثل في الانتقال من بيان الى آخر بسرعة الأمر الذي لا يساعده كثيراً في حفظ أرقام الدليل .

والفصل في هذا النقاش يتوقف على مدى دقة عملية المراجعة، فإذا كانت المراجعة قد اجريت بدقة ولنا ثقة كبيرة في القائمين بها فإن الطريقة الاولى تكون أفضل حيث انها تسهل العمل كثيراً على الموظفين والعكس صحيح

٨ – توجيه نظر الموظفين الى ضرورة الربط بين مختلف البيانات الخاصة بوحدة واحـــدة اذا كان هناك غموض يكتنف بعض البيانات ، فقد يؤدي الربط بين البيانات الى كشف هذا الغموض .

٩ – توجيه نظر الموظفين الى الحالات التي يوقع فيها دليل « لم يبين » وهي الحالات التي يعتورها الغموض ويجب محاولة الوصول الى تبيانها باتباع ما ورد في الفقرة السابقة قبل ان يقرر الموظفون اعطاءها دليل « لم يبين » ويحسن في مثل هذه الحالات استشارة المشرف على عملية الترميز .

١٠ – لفت نظر الموظفين الى الاخطاء التي يمكن ان يقعوا فيها اذا لم ينتبهوا جيداً أثناء عملهم .

١١ – لفت نظر الموظفين الى الاخطاء التي يحتمل وجودها في البيانات والتي يجب عليهم تصحيحها قبل الاستمرار في عملية الترميز .

١٢ – تحديد الطريقـــة التي يمكن بواسطتها ان يستفسر الموظفون عن الصعوبات والمشاكل التي تواجههم في العمل على ان يكون التبليغ عنها كتابة.

١٣ – توجيه نظر الموظفين الى متابعة التعليمات التي تظهر أولاً بأول وان يحتفظوا بنسخة منها لديهم .

١٤ – تحديد المسائل التنظيمية المختلفة مثل : –

أ – كيفية استلام الاستمارات .

ب – كيفية اعادتها بعد ترميزها .

ج – كيفية اعداد التقارير عن سيرالعمل بالنسبة لكل موظف.

د – مواعيد تقديم هذه التقارير .

هـ – المعدل اليومي لانتاج الموظفين .

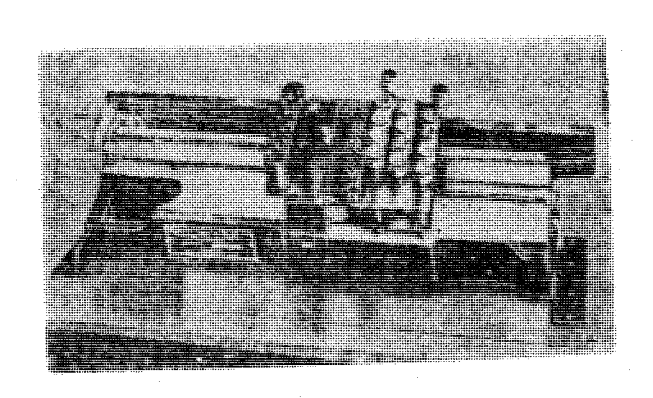
١٥ – تقديم مثال لبيانات على استمارة وترميزها بكل دقة وعناية .

١٦ - النص على أن يوقع المرمزون ومراجعو الترميز على الاستمارات التي ينتهون منها وان يكون التوقيع مصحوبا بالتاريخ ومكان ذلك التوقيع .

١٧ - النص على تقسيم العمل بين الموظفين اذا كان هناك رغبة في ذلك ، حيث يفضل البعض أن يقوم بعض الموظفين بترميز البيانات الصعبة مثل المهنة والنشاط الاقتصادي بينما يقوم الموظفون الباقون بترميز البيانات الاخرى . وعند اتباع هذه الطريقة يراعى في الفريق الذي يقوم بتوقيع الدليل الصعب أن يكون أكثر مرونة ودقة في العمل . على أن اتباع هذه الطريقة يؤدي الى تضخيم حجم العمل والقسم القائم به الى حد ما الا انه في نفس الوقت ادعى الى الدقة والسلامة اذا أمكن التغلب على العيب الناشيء من ضخامة العمل

التثقيب ومراجعته :

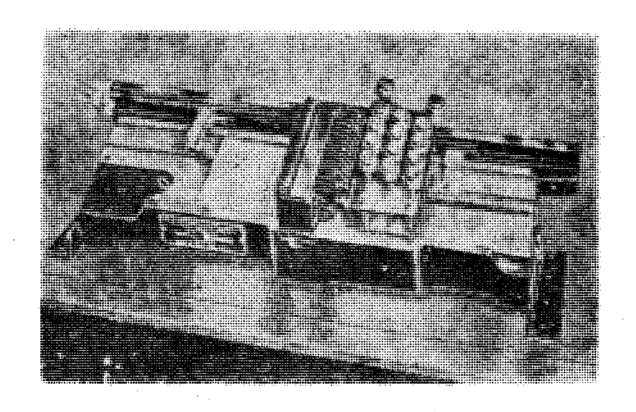
والخطوة التالية هي تحويل الرموز التي كتبت أمام الاجابات المختلفة إلى ثقوب في مواضع معينة من بطاقة التثقيب . ولاجراء ذلك يخصص لكل سؤال (لكل بيان) عمود أو اكثر في البطاقة بحسب عدد الأرقام التي يتكون منها آخر عدد في دليل هذا البيان . فاذا كان دليل البيان مكون من رقم واحد ، أي إذا كانت الاجابات المحتملة للسؤال لا تصل الى العدد ١٠ خصص للسؤال عموداً واحدا في البطاقة ، وإذا كان دليل اجابات السؤال يتكون من رقمين أي أن الاجابات المحتملة لهذا السؤال يمكن ان تصل الى ٩٩ في الحد الأقصى خصصنا للسؤال عمودين احدهما لرقم **الآحاد** والثاني لرقم العشرات . وتجري عملية التثقيب بعمل ثقب (آليا) في الأعمدة المخصصة لاجابات السؤال وأمام الأرقام الخاصة باجابة معينة في الاستمارة ، مثلا اذا كان رقم الدليل لاجابة معينة هو ٢٥ يعمل ثقب أمام رقم ٢ في عمود العشرات وثقب أمام رقم ٥ في عمود الآحاد .



آلة التثقيب اليدوية

وتجري عملية التثقيب بواسطة آله تسمى بآلة التثقيب وهي تحتوي على عشرة أزرار مكتوب على كل منها احـــد الأرقام من صفر الى ٩ . وعند الضغط على احد الأزرار وليكن الزر رقم ٨ تقوم الآلة بعمل ثقب في مكان الرقم ٨ في البطاقة . وعند بدء العمل توضع البطاقة في داخل الآلة ، وبذلك تكون الآلة مستعدة للتثقيب في العمود الاول من البطـــاقة ، فإذا ضغطنا على الزر الخاص برقم معين فتثقبه الآلة في هــذا العمود وتنتقل البطـاقة آلياً الى العمود الثاني أي تندفع البطاقة آلياً إلى الخارج مسافة مساويــة للمسافة الواقعة بـــين عمودين من أعمدتها (وهــذه الحركة مماثلة لحركة قطعة الورق في حالة الآلات الكاتبة) ، ويستمر الثاقب في الضغط على الازرار حسب ارقام الدليل المدونة امـام الاجابات المختلفة حسب ترتيبها في اعمدة بطاقة التثقيب حتى ينتهي منهـا جميعاً . وعندئذ يضغط على زر خاص فتخرج البطاقة ويسحبها الثاقب خارج الآلة ثم يضع بطاقة ثانية للوحدة الثانيةويشرع في تثقبها وهكذا. وبذلك نلاحظ ان كمية البيانات التي نثقبها في بطاقة التثيب تتوقف على كمية البيانات التي نريد تثقيبها وعلى عدد الاعمدة التي يحتاجها دليل هذه البيانات وعلى عــدد

الاعمدة الموجود فعلا في بطاقة التثقيب ، فاذا كانت البيانات الخاصة بكل وحدة من وحدات الدراسة يحتاج تثقيبها الى ٧٩ عمود وكانت اعمدة البطاقة هي ٦٥ فقط نحتاج بذلك الى بطاقتين لتثقيب البيانات الخاصة بوحدة واحدة ، اما اذا كانت اعمـــدة البطاقة ٨٥ عمودا نحتاج بذلك الى بطاقة واحدة لبيانات كل وحدة . لهذا نلاحظ ان كثرة اعمدة البطاقة تعتبر ميزة كبيرة في بطاقات التثقيب ، والموجود في الوقت الحاضر من هذه البطاقات يحتوي اما على ٣٦ عمود (وهو صنف قديم جداً) لم يعد يستعمل كثيرا ، أو ٦٥ عمود أو ٨٥ عمود . وبالطبع كل نوع من آلالات الاحصائية يكون معدا له نوع معين من بطاقات التثيب (من حيث عدد الاعمدة) ، ولذلك عندما نتكلم عن آلالات الاحصائية نقول آلات ٣٦ عمود او آلات ٦٥ عمود او آلات ٨٥ عمود .



آلة المراجعة اليدوية

وحيث انه يمكن ان تحدث اخطاء عند اجراء عملية التثقيب ، اذ قـد يضغط الثاقب على الزر الخاص برقم ٨ بينما يكون رقم الدليل ٧ ، لهـذا

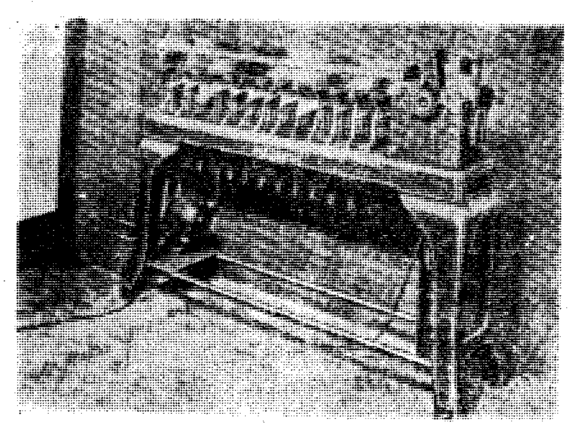
يكون من الواجب مراجعة هذه العملية . وتتم عملية المراجعة آليا كذلك بواسطه آلة خاصة بذلك لا تختلف كثيرا عن آلة التثقيب السابقة . ويقوم مراجع التثقيب بادخال البطاقة المثقبة في آلة المراجعة ويضغط على الازرار وفقا لارقام الدليل الخاصة بالاجابات المختلفة وتبعا لترتيبها حسب اعمدة البطاقة ، فاذا كان الرقم الذي يضغط عليه يطابق الرقم المثقوب في البطاقة تندفع البطاقة الى الخارج ويدل ذلك على صحة التثقيب وبذلك يستمر المراجع في عمله ، أما اذا كان الرقم الذي يضغط عليه لا يطابق الرقم المثقوب في البطاقة لا تتحرك وحينئذ ينتبه المراجع الى وجود خطأ في التثقيب ، وحيث انه يمكن أن يكون هو الذي أخطأ لذلك يجب أن يتأكد من صحة ضغطه على الزر فاذا كان الأمر كذلك علم أن هناك خطأ في عمل المثقب وبذلك يسحب البطاقة من الآلة حتى تمزق وتعمل بطاقة اخرى بدلها .

التصنيف (الفرز) :

باجراء الخطوات السابقة نكون قد حولنا البيانات الموجودة في الاستمارة الى ثقوب في أعمدة البطاقات طبقاً لارقام الدليل الخاصة بها ، وبذلك يتجمع لدينا بطاقات مثقبة بحيث يكون لكل وحدة استبيان بطاقة أو أكثر حسب نوع البطاقة التي نستخدمها وكمية المعلومات التي ثقبناها لكل وحدة، والأجراء التالي هو تصنيف هذه البطاقات أي فرزها بالنسبة لكل بيان من البيانات التي ثقبناها فاذا كان لدينا بطاقات تمثل المعلومات المختلفة عن السكان وأردنا مثلاً تصنيف السكان تبعاً للجنس ، لذلك نعرف أولا العمود المخصص لتثقيب هذا البيان وليكن العمود رقم ٥، ثم نمرر البطاقات في الات كهربائيه تسمى بالآلات الفرز حيث نضبط مؤشر الآلة على العمود رقم ٥ (يوجد على الآلة مقياس مدرج مكتوب عليه أرقام الأعمدة الموجودة في البطاقة ، أي الأرقام ١، ٢،... الى رقم آخر عمود في البطاقة وليكن ٨٥ ، ويتحرك أمام هذا المقياس مؤشر بحيث اذا أوقفنا هذ المؤشر أمام رقم معين على المقياس

المدرج فان ذلك يعني ان الآلة ستقوم بتصنيف أي فرز البطاقات حسب الأرقام المثقبة في العمود الذي يحمل هذا الرقم في البطاقة) . ولذلك بضبطنا المؤشر أمام رقم ٥ تكون الآلة مستعدة لفرز البطاقات حسب الأرقام المثقبة في هذا العمود فاذا كان دليل الجنس هو ١ للذكور و٢ للاناث، فان ادارتنا للالة تجعلها تدفع البطاقات دفعاً سريعاً داخل مجرى معين وأثناء اندفاعها تسقط البطاقات المثقبة عند الرقم ١ في العمود الخامس فقط في جيب خاص بالرقم ١ ، أما البطاقات المثقبة عند الرقم ٢ في العمود الخامس فتسقط في جيب خاص بالرقم ٢ أما اذا كانت البطاقات ليس بها أي ثقب (بسبب نسيانها عند التثقيب) فتسقط في جيب خاص بذلك .

بعد انتهاء فرز البطاقات تبعاً للعمود الخاص بالجنس تعطينا الآلة حسب عدادها الخاص عدد البطاقات ذات الرقم ١ وهم الذكور وذات الرقم ٢ وهم الاناث . يمكننا بعد ذلك أن نعيد وضع البطاقات في الآلة لفرز البيان الخاص بالحالة الزواجية ، فاذا كان هذا البيان مثقب في العمود رقم ٦ نضبط المؤشر أمام هذا الرقم وندير الآله فتندفع البطاقات في مجراها الخاص وتسقط بعضها في الجيب رقم ١ وبعضها في الرقم ٢ وبعضها في الرقم ٣ وبعضها في الرقم ٤ وبعضها في الجيب الأخير . وبذلك يعطينا عداد الآلة عدد البطاقات ذات الرقم ١ أي السكان العزاب ، والبطاقات رقم ٢ أي المتزوجين ، والبطاقات رقم ٣ أي المطلقين ، والبطاقات رقم ٤ أي الأرامل والبطاقات الأخيره هم السكان دون سن الزواج ، وذلك اذا كانت هذه الآرقام هي الدليل الرقمي للحالة الزواجية .



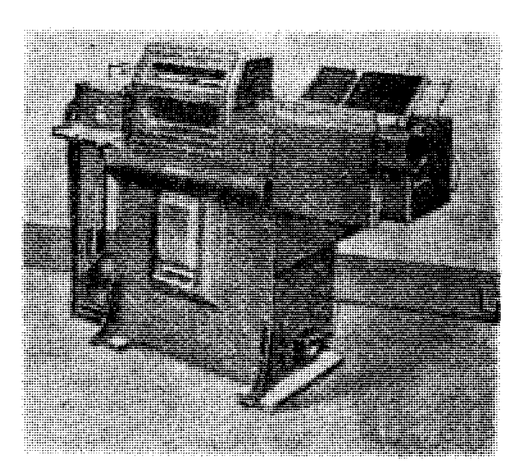
آلة الفرز الكهربائية

التبويب :

اذا كانت آلة الفرز تقوم بتصنيف البطاقات وعدها في كل فئة نريدها فما هي وظيفة آلة التبويب ؟ نفترض اننا فرزنا البطاقات الخاصة بالمشتغلين وبذلك عرفنا عدد الذكور منهم وعدد الاناث ، ونريد بعد ذلك أن نعرف جملة الاجور التي تقاضاها كل من الجنسين ، نمرر البطاقات بعد فرزها داخل آلة كهربائية تسمى آلة التبويب . وهي تقوم بقراءة الأرقام المثقبة في الأعمدة الخاصة بالاجور في البطاقات ثم تحولها الى أرقام ثم تجمع هذه الأرقام مع بعضها بطاقة بعد اخرى وتسجل المجموع على قطعة ورق مثبتة في الآلة ويمكن نزعها منها . بهذه الطريقة نستطيع أن نعرف جملة الاجور التي تقاضاها كل من الجنسين للمشتغلين مطبوعة ومبوبة على ورقة خاصة .

وآلة التبويب تستخدم لتحقيق اغراض مختلفة ، واستخدامها مبنى على فكرة انها تستطيع أن تقرأ الثقوب الموجودة في أي عمود في البطاقة المثقبة ثم تحولها إلى الارقام الخاصة بها ثم تطبعها على ورق خاص ، فاذا كنا نريد أن نحصل على بيان عن العمر والاجر لكل مشتغل تبعـــاً للجنس ، نقوم بتصنيف البطاقات تبعاً للجنس في آلة الفرز ثم نمررها في آلة التبويب فتقوم

(طبقا لنظام خاص) بتسجيل البيان الخاص بالعمر والاجر امام كل عامل في كل جنس من الجنسين ، أي انها تطبع لكل جنس عمودين احدهما يحتوي على العمر والآخر يحتوي على الأجر وبذلك تحصل على ورقة عليها عمر واجر كل عامل من الجنسين . ولذلك يمكن أن تستخدم هذه الآلات في أعداد القوائم الخاصة بمرتبات الموظفين في أي مؤسسة يعمل بها عدد كبير جداً من المشتغلين . وتستطيع آلة التنويب معالجة ٣٥٠٠ بطاقة في الساعة كما انها تستطيع تسجيل سبعة مجاميع مختلفة في نفس الوقت ، ولذلك فهي تسهل علينا العمل كثيرا وتساعدنا في الحصول على النتائج بسرعة فائقة .



آلة التبويب الكهربائية

عرض البيانات في جداول :

بعد جميع المعلومات وتصنيفها سواء باليد أو استخدام الآلات الاحصائية تكون الخطوة التالية في سلسلة العمليات الاحصائية هي عرض هذه المعلومات . وهناك طريقتان للعرض يمكن استعمالهما هما الجداول الاحصائية والرسوم البيانية .

والجدول الاحصائي يختلف عن أي كشف آخر،فالكشف يظهر معلومات

فردية أي عن كل وحدة على حدة بينما يظهر الجدول الاحصائي المعلومات بعد تجميعها وتصنيفهـــا في فئات معينة قد تكون فئات جغرافية أو زمنية أو وصفية أو كمية ، وقد تكون تهجينا لهذه الأنواع بحيث تظهر المعلومات في الجدول مصنفة تبعاً لنوعين من هذه الفئات أو أكثر . وبذلك عند عرض البيانات في الجداول الاحصائية تضيع معالمهـــا الشخصية وتصبح مجرد رقم معين في فئة معينة .

وبذلك يكون الجدول الاحصائي هو حلقة الاتصال بين الهيئة الاحصائية التي أعدته وبين كل من سوف يطلع عليه في المستقبل لأخذ المعلومات الـــتي يريدها، لذلك يجب أن يفسر نفسه بنفسه فلا يحتاج بذلك الى أي استفسار . لهذا لا بد أن تتوفر في الجدول الاحصائي شروط معينة حتى يكون واضحاً لا يشوبه أي ابهام أو أي شك في محتوياته .

عناصر الجدول الاحصائي وشروطه :

عند تصميم أي جدول احصائي يجب مراعاة القواعد الآتية :

أولاً – العنوان :

يجب أن يكون للجدول عنوان يجيب على الأسئلة الآتية :

أ – ما هي الأرقام التي يحتوي عليها الجدول ؟

ب – أي مكان أو أماكن تتعلق بها هذه الارقام ؟

ج – أي فترة زمنية تتعلق بها هذه الأرقام ؟

د– ما هو الأساس أو الأسس التي صنفت تبعاً لهاالارقام في هذا الجدول؟

مثـــال

الخيارات الزراعية في لبنان موزعة تبعاً لموقعها عام ١٩٦٥ .

وعند اعداد عنوان الجدول يجب مراعاة ان يظهر في اعلا الجدول ، واذا كان طويلا يحسن كتابته في شكل هرم مقلوب ، كما يجب ان يكون دقيقاً فيما يختص بالارقام الموجودة فيه :

مثـــال

الوفيات المسجلة في مكاتب وزارة الصحة في سوريا موزعة تبعاً لمناطق الدولة وسبب الوفاة عام ١٩٦٥

ونلاحظ اننا كتبنا العنوان لأنه طويل في شكل هرم مقلوب ، واضفنا الفقرة -المسجلة في مكاتب وزارة الصحة- حتى لا تؤخذ الارقام على انها تمثل جميع الوفيات في الدولة . واذا كانت الارقام التي يظهرها الجدول هى ارقام قياسية يجب أن يظهر تحت العنوان فترة الاساس لهذه الارقام :

مثـــال

الارقام القياسية لنفقة المعيشة في
الجمهورية العربية المتحدة
(عام ١٩٣٥ = ١٠٠)

ثانيا - عناوين الاعمدة :

كل عمود في الجدول يجب ان يكون له عنوان يوضح الارقام والبيانات المدونة فية توضيحا كاملا وبايجاز . واذا لا حظنا ان تركيز عنوان أي عمود في كلمات قليلة يجعله غير واضح يجب توضيحه توضيحاً كاملاً في ملاحظة خاصة أسفل الجدول ، مثلا قد يكون عنوان العمود عدد السكان ولكننا

نريد ان نوضح ان هذه الاعداد لا تشتمل على البدو والسكان الذين وجدوا خارج الدولة وخارج مياهها الاقليمية وقت اجراء التعداد . ولكننا لا نستطيع أن نكتب كل هذا التوضيح في عنوان العمود ، لذلك نضطر الى تدوينه في اسفل الجدول .

ثالثا : عناوين الاسطر في كعب الجدول .

التوضيح الوصفي للارقام في كل سطر يكتب في العمود الاول من الجهة اليمنى من الجدول الذي يسمى بكعب الجدول . وعند تنظيم كعب الجدول يجب ملاحظة أن نكتب في قمة الكعب البنود الاكثر اهمية ، واذا كانت هذه البنود هي فترات زمنية يجب أن نبدأ بالفترات الاقدم ثم نتسلسل زمنياً حتى نصل الى أحدث فترة يظهرها الجدول . كذلك اذا كانت البيانات في كعب الجدول هي سنوات يحسن تقسيمها الى مجموعات خمسية يفصلها عن بعضها فراغات حتى تسهل قراءتها وحتى تبدو في مظهر أنيق . أما اذا كانت شهور يكون من الافضل تقسيمها الىمجموعات تتكون كل مجموعة منها من ثلاث شهور . واذا كان التوضيح في كعب الجدول يقسم الارقام الى مجموعات رئيسية واخرى فرعيةتابعة لها يجب ترك بعض الفراغ تحت المجموعة الرئيسية ابتداء من اليمين قبل كتابة المجموعات الفرعية التابعة لها . كذلك اذا كان عنوان الارقام في سطر معين في الجدول يحتاج الى أكثر من سطر واحد تكتب الارقام أمام السطر الاخير في العنوان وليس امام الأسطر الأولى .

رابعاً – المجاميع الرئيسية والفرعية :

يجب أن يظهر الجدول المجاميع الخاصة بالأرقام الموجودة فيه كلما كان ذلك أمراً ممكناً سواء كان ذلك خاص بمجاميع أرقام الاعمدة أو ارقام

الأسطر . وإذا احتوى الجدول على مجاميع رئيسية وأخرى فرعية يجب كتابتها في الجدول بطريقة تساعد على الاهتداء اليها والتمييز بينها بسهولة . ويلاحظ أن تكتب المجاميع في نهاية الأرقام المكونة لها الا اذا كان المطلوب أن يعطى الجدول الأهمية الاولى للمجاميع وبذلك يمكن كتابتها في قمة الأرقام المكونة لها . كذلك يجب التحقق من صحة المجاميع حيث يكون أمراً مخزياً أن يجد من يطلع على الجدول خطأ فيها اذ أن ذلك يجعله يفقد الثقة في جميع الارقام الواردة في الجدول حيث انه لن يستطيع أن يتبين مصدر الخطأ . واذا كان هناك تقريب في الارقام بحيث يكون هناك بعض الفرق بينها وبين مجاميعها يجب الاشارة الى ذلك تحت الجدول في ملاحظة خاصة .

خامساً – المتوسطات والنسب المئوية :

اذا احتوى الجدول على المجاميع والمتوسطات يجب أن تسبق المجاميع المتوسطات . واذا كانت الارقام في عمود ما في الجدول عباره عن نسب مئوية يجب أن نشير في عنوان العمود الى نوع الرقم المنسوب اليه وبذلك يجب تجنب ذكر العلامة ٪ فقط اذ يجب أن نذكر ٪ من المجموع أو ٪ من الزيادة او ٪ من النقص . . وهكذا . كذلك يجب أن يكون مجموع النسب ١٠٠ ٪ اذا كانت النسب حسبت من المجموع ولكن قد يكون المجموع الفعلي للنسب يزيد أو ينقص ١٠٠ بسبب التقريب ، وفي هذه الحالة يكتب المجموع ١٠٠ ٪ بحيث لا يزيد المجموع الفعلي للأرقام عن ١٠٠٫٢ ولا ينقص عن ٩٩٫٨ ، فاذا زاد الفرق عن ٠٫٢ ٪ يجب مراجعة العمليات الحسابية الخاصة بالنسب المئوية حتى نتأكد من عدم وجود أخطاء .

سادساً – الوحدات :

اذا كانت الارقـــام في جميع الأعمدة مقاسة بنفس الوحدة يكتب نوع الوحدة بين قوسين تحت العنوان العام للجدول . أما اذا اختلفت وحدات

الارقام من عمود الى آخر أو من سطر الى آخر يكتب نوع الوحدة الخاصة بالعمود مع عنوانه والوحدة الخاصة بكل سطر مع عنوانه في كعب الجدول. واذا اختلفت وحدات الارقام في عمود معين من سطر الى آخر يحسن أن يظهر الجدول عموداً خاصاً مجاوراً تظهر فيه وحدات الارقام كل على حدة (يظهر ذلك في جداول التجارة الخارجيه الخاصة بكميات البضائع حيث تكون وحدة الكمية لكل نوع مختلفة عن الاخرى) .

سابعاً - تدوير الارقام :

اذا كانت جميع الأرقام في الجدول مدورة أي محذوف منها نفس العدد من الارقام يكتب ذلك في العنوان العام للجدول، فاذا كانت جميع الارقام مختصرة على مليون مثلا يكتب ذلك في العنوان العام للجدول . واذا كانت الارقام في عمود معين فقط محذوف منها نفس العدد من الارقام يكتب ذلك في عنوان العمود نفسه (مثلاً الف دينار) . واذا كانب الارقام في سطر معين محذوف منها نفس العدد من الارقام بكتب ذلك في التوضيح الخاص بالسطر في كعب الجدول .

ثامناً - الملاحظات :

تكتب الملاحظات التوضيحية للارقام أو العناوين في الجدول في أسفله . فاذا كانت هذه الملاحظات عامة بالنسبة لجميع الارقام في الجدول تكتب في أسفله دون أية اشارة . اما اذا كانت خاصة ببعض الارقام أو ببعض العناوين يجب وضع اشارة معينة على الرقم أو العنوان وكتابة نفس الاشارة أمام الملاحظات الخاصة بذلك اسفل الجدول، واذا تعددت الملاحظات يجب استعمال اشارات مختلفة ، وغالباً تستعمل الارقام كاشارات مع العناوين والأشكال الهندسية كاشارات مع الارقام . وبشكل عام لا يجب أن يترك أي عنوان

أو أي رقم غامضاً أو به بعض الشك إذ يجب دائماً التوضيح والتفسير في ملاحظات أسفل الجدول .

وهناك اشارات اتفق عليها من قبل الهيئات الدولية يمكن ان تستعمل للدلالة على بعض المفاهيم مثل :

. . . . وتدل على ارقام غير متوفرة .

ــــــ وتدل على ارقام غير موجودة أصلا

× وتدل على ارقام غير رسمية .

[ ] وتدل على ارقام غير محسوبة ضمن المجموع .

وبالرغم من أن هناك اتفاق عام على معنى هذه الاشارات ، الا انه يجب عدم اغفال ذكر دلالتها في مقدمة النشرة .

تاسعاً : المصدر

يجب الاشارة إلى مصدر الارقام في اسفل الجدول تحت الملاحظات اذا لم يكن مصمم الجدول المسئول عنها اصلا . وعند ذكر المصدر يجب كتابه اسم المؤلف او الهيئة المسئولة اولا ثم عنوان الكتاب أو النشرة أو المقال أو المجلة ثم أسم الناشر ثم تاريخ النشر ثم رقم الصحيفة الموجودة بها الارقام ويستحسن وضع عنوان المصدر بين قوسين من اعلا .

عاشرا : الخطوط الافقية التي يحتوي عليها الجدولى .

اذا كانت المجاميع واردة في اول الجدول او ضمن جسمه يحسن فصلها عن الارقام التي تأتي بعدها بخطين متجاورين ، ولا يجب أن تقطع هذه الخطوط كعب الجدول ، وفيما عدا ذلك لا يجب أن يظهر في الجدول من الخطوط الافقية غير الخطوط الخاصة بعناوين الاعمدة وبالمجاميع في نهاية الجدول من أسفل .

ملاحظات عامة :

١ – يحسن ترقيم الجداول التي تحتوي عليها النشرة الاحصائية حتى يمكن الاهتداء اليها بعد الاطلاع على فهرس الجدوال في مستهل النشرة .

٢ – اذا كانت جميع الجداول في النشرة خاصة بنفس الدولة وبنفس الفترة الزمنية لا يجب أن نذكر ذلك في عنوان كل جدول اذ يكفي كتابتها مع العنوان العام للنشرة . ويجب أن لا ننسى اسم منطقة معينة أو فترة زمنية معينة في عنوان جدول معين اذا لم تكن الأرقام في هذا الجدول خاصة بالدولة عامة أو بالفترة الزمنية العامة للنشرة .

٣ – يمكن ترقيم عناوين الأعمدة وعناوين الأسطر في جدول معين حتى يسهل شرح محتوياتها ويكون ذلك أمرا واجبا في التقارير الاحصائية التي تحتوي على جداول وشرح للأرقام الواردة فيها .

٤ – يجب أن يكون الشكل العام للجدول مقبولاً فلا يجب أن يكون طويلاً وضيقاً او قصيراً جداً وواسعاً كما لا يجب أن تظهر الأرقام فيه متضاربة ومعقدة المنظر ، الأمر الذي ينفر القاريء . ويجب أن تظهر الأرقام في كل عمود مرتبة تحت بعضها من ناحية الخانات المختلفة التي تتكون منها .

التوزيع التكراري :

من أهم الخدمات التي يقدمهـــا الاحصاء للابحاث المختلفة كيفية تنظيم واختصار البيانات بشكل يسمح للعقل أن يتفهمها ويقف على أهم خصائصها . ومن أهم الوسائل التي يستخدمها الاحصائيون لهذا الغرض هو عمـــل توزيع تكراري لتلك البيانات . فلو فرض أن كان لدينا قيماً كثيرة لأطوال مجموعة من الأفراد وليكن عددهم الفان مثلاً فليس من السهل على العقل أن يستوعب مميزات طول هؤلاء دون تلخيص وتنظيم مبدئي .

فيصعب على القارىء أن يعرف أصغر القيم وأكبرها لأطوال هؤلاء دون ضياع بعض الوقت والجهد ، وبالمثل يصعب معرفة ما اذا كان هناك تركز في الأطوال عند قيم معينة بمجرد النظر الى القيم الدالة على أطوال هؤلاء الاشخاص، وكذلك يصعب بمجرد النظر الى القيم الاصلية معرفة ما اذ كان هناك استمرار في التغير أي معرفة ما اذا كان هناك أشخاص يمثلون جميع الاطوال بين حدي المدى .

اذا فرض ان كانت البيانات الاصلية محدودة العدد - حوالي العشرة أو العشرين قيمة مثلاً - فمن الجائز أن نضعها في شكل مرتب اما في وضع تصاعدي أو تنازلي. ويطلق على القيم في وضعها الجديد اسم التوزيع المنتظم.

فاذا أعطينا القيم الآتية وهي تدل على اجور عشرة من العمال في اليوم : ١٥ ، ١٠ ، ٤٠ ، ١٩ ، ١٨ ، ٣٠ ، ١٦ ، ٢٥ ، ٢٠ ، ٢٠ ، فقد يحسن ترتيب تلك الأجور ترتيباً تصاعدياً لكي يقف القارىء أولا على مدى تغير القيم المعطاة (Range) وثانياً تركزها عند قيمة أو قيم معينة أو انعدام هذا التركز (Concentration) وثالثاً على استمرار هذه القيم خلال المدى كله (Continuity) أو انعدام هذا الاستمرار في بعض أجزائه وستكون القيم بعد ترتيبها هكذا .

١٠ ، ١٥ ، ١٦ ، ١٨ ، ١٩ ، ٢٠ ، ٢٠ ، ٢٥ ، ٣٠ ، ٤٠

فواضح ان مدى التغير في الاجور لهؤلاء العمال يتراوح بين ١٠ — ٤٠ ليرة في اليوم وأن عدداً لا بأس به من أفراد المجموعة يتركز أجره عند ١٥ ليرة الى ٢٠ ليرة وأنه ليس هناك قيم بين ٣٠ - ٤٠ ليرة وهي نهاية المدى .

ومن الواضح انه لم يكن من السهل الوصول الى تلك الحقائق دون استخدام فكرة التوزيع المنتظم . ولكن لا يمكن الاكتفاء بهذه الفكرة كوسيلة لاختصار البيانات إذا زاد عددها زيادة كبيرة فلا بد في مثل تلك الحالات من الالتجاء الى التوزيع التكراري .

والفكرة الأساسية في التوزيع التكراري هي تحديد عدد مرات ظهور أي قيمة معينة ثم تسجيل تلك القيم في جدول وأمام كل منها عدد مرات ظهورها . ولتوضيح تلك الفكرة عملياً نفترض انه طلب الينا اختصار البيانات الآتية وهي عبارة عن أجور ٢٥٠ عامل في مصنع ما :

١٦ ٢٠ ٨ ١٢ ١٣ ٢٨ ٤٣ ٣٤ ٢٣ ٢٠ ٢٤ ٤٥ ٣٦ ٢٩ ٣٥ ٢٥ ٢٠ ١٤ ١١ ١٠

١٦ ١٢ ٤٠ ٣٢ ٢٤ ٢٠ ١٨ ١٥ ٨ ٢٢ ١٨ ٢١ ٢٦ ٣١ ٣١ ١٨ ٢١ ٢٦ ١٥ ٢٧

٢٠ ١٥ ١٧ ٢١ ١٥ ١٨ ١٩ ٦٩ ٢٥ ٢٣ ٣٩ ١٩ ١٠ ٧ ٤ ٣٧ ٥٠ ١٦ ٣٣ ٣٥

٢٧ ١٦ ٩ ٦ ٤٥ ٣٨ ١٥ ١٧ ١٧ ٣٠ ٢٧ ٣٦ ٤٥ ١٢ ١٦ ١٨ ٢٠ ٢٢ ٢٤ ٢٧

٣٣ ٣٧ ٤٢ ٢٢ ١٨ ١٥ ٢٧ ٢٢ ١٨ ١٥ ١١ ١٧ ١٩ ١٤ ٢١ ٣٠ ٢٥ ٢١ ١٩ ١٦

٢٠ ١٩ ١٧ ١٤ ١٨ ٢٥ ٣٠ ٣٥ ٢٣ ١٥ ١٥ ٢١ ١٠ ٢٥ ١٣ ٣٠ ١٨ ٣٨ ٣٠ ٣٦

٢٥ ٢٩ ٣٤ ١٧ ٢٢ ٢٥ ٢٧ ٢٨ ٣٥ ٣٠ ٢٥ ١٨ ٢٩ ١٦ ٣٨ ١٠ ١٦ ٢٠ ١٣ ٢٣

٣٦ ٣٢ ٢٠ ١٦ ٢٠ ١٨ ٥ ٢٧ ٣٢ ٣٨ ١١ ٢٠ ١٩ ٢٥ ٢١ ٣٠ ٢٦ ٢٠ ١٥ ١٩

١٤ ٢٥ ١٣ ١٦ ١٧ ١٩ ٢١ ٢٣ ١٥ ٢٥ ١٧ ٣٥ ٢١ ١٠ ١٩ ٢٥ ١٥ ١٧ ١٧ ٢٠

١٩ ٢٩ ٣٥ ٢٧ ٢٤ ٩ ١٧ ١٩ ٢٥ ٣٥ ٣٣ ٢٨ ٢٠ ١٢ ٨ ٥ ١٥ ٢١ ٣٥ ٣٥

٣٠ ٣٥ ٢٥ ٢٦ ٢٠ ١٩ ٣١ ١٧ ١٤ ١٥

٢٢ ٢٢ ٢٢ ١٨ ١٢ ٣٠ ٣٠ ١٩ ٢١ ٥

١٩ ١٩ ٢٠ ٢٣ ١٦ ١٤ ١٠ ٣٠ ٢٥ ٢١

١٧ ١٥ ١٩ ٢٣ ٢٥ ١٨ ٤٠ ٤٠ ٤١ ٢٤

١٥ ٢٩ ٣٤ ٢٨ ٢٧ ١٥ ٢٨ ٣٠ ٢٥ ١٩

وطريقة عمل التوزيع التكراري وفقاً لتعريفه هي تسجيل قيم المتغير واحدة بعد أخرى حسب ورودها في الكشف الأصلي ثم وضع علامات امام

كل قيمة يتكرر حدوثها وذلك لتحديد عدد مرات تكرارها ويسهل بعد ذلك ترتيب القيم ترتيبا تصاعديا أو تنازلياً وأمام كل قيمة تكرارها .

والواقـــع ان هذه الطريقة المبدئية في عمل التوزيع التكراري لا تفيدنا كثيراً إذ ان تكرارات القـــيم المختلفة ستتزايد وتتناقص بطبيعة الحال بلا نظام ولذا يجب ضم بعض القيم الى البعض الآخر وتكوين فئـــات (groups) أو (Classes) فلو أخذنا العمال الذين يتقاضون أجور تتراوح بين ٤، ٨ ليرة لحصلنا على تكرار قدره ٦ لهذه الفئة ، وبالمثل لو أخذنا الذين يتقاضون بين ٨ ، ١٢ لحصلنا على تكرار قدره ١٤ لتلك الفئة . وهكذا اذا كررنا عملية ضم القـــيم وعد تكراراتها لكان الناتج جدولا تكرازيا أنسب في صورته لتحقيق الغرض الذي من أجله قمنا بعمل التوزيع التكراري .

فئات الاجور بالليرة	التكرار (عدد العمال)
٤ —	٦
٨ —	١٤
١٢ —	٣١
١٦ —	٥٤
٢٠ —	٤٣
٢٤ —	٣٦
٢٨ —	٢٧
٣٢ —	١٩
٣٦ —	١٠
٤٠ —	٦
٤٤ —	٣
٤٨ — ٥٢	١

ويستطيع القارىء أن يحس بنفسه فائدة التوزيع التكراري في اختصار البيانات فقد كان يستحيل معرفة مدى تغير الأجور دون بذل مجهود ووقت في حين ان ذلك واضح لأول وهلة من مجرد قراءة التوزيع التكراري .

وبالمثل لا يتضح تركيز القيم عند أي قيمة معيّنة او انعدام هذا التركز لمن يكتفي في دراسته بالكشف الاصلي ولكن على العكس من ذلك واضح جلي ان الاجور تتركز عند الأجر من ١٦ الى ٢٠ ليرة .

وكذلك لا يمكن الجزم باستمرار التغير في الأجور في المدى كله او انعدام هذا الاستمرار لمن يطلع على كشف الأجور في حين انه من أيسر الأمور الحكم على هذا بالنظر في جدول التوزيع التكراري حيث يتضح استمرار التغيير بين ٤٠ ليرة ، ٥٠ ليرة . تلك الحقائق الثلاث ، تحديد المدى والتركز او انعدامه والحكم على استمرار التغير خلال المدى أو عدمه من أهم الفوائد العملية للتوزيعات التكرارية .

وليس من الضروري عند عمل التوزيع التكراري أن يحدد عدد مرات تكرار كل قيمة على حدة ثم ضم القيم في فئات كما ذكرنا اذ كان الغرض من وصف تلك الطريقة تأكيد الفكرة الأساسية التي ينبنى عليها أي توزيع . أما الطريقة العملية لذلك فهي تحديد مدى التغير للظاهره المراد وصفها ثم تقسيم هذا المدى الى فئات يستحسن أن تكون متساوية في الطول ووضعها في جدول تفريغ كما هو موضح في الجدول الآتي .

جدول تفريغ أجور ٢٥٠ عامل .

فئات الاجور	التفريغ	التكرار (عدد العمال)
٤ —	𝍸 𝍷	٦
٨ —	𝍸 𝍸 𝍷𝍷𝍷𝍷	١٤
١٢ —	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍷	٣١
١٦ —	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍷𝍷𝍷𝍷	٥٤
٢٠ —	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍷𝍷𝍷	٤٣
٢٤ —	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍷	٣٦
٢٨ —	𝍸 𝍸 𝍸 𝍸 𝍸 𝍷𝍷	٢٧
٣٢ —	𝍸 𝍸 𝍸 𝍷𝍷𝍷𝍷	١٩
٣٦ —	𝍸 𝍸	١٠
٤٠ —	𝍸 𝍷	٦
٤٤ —	𝍷𝍷𝍷	٣
٤٨—٥٣	𝍷	١
المجموع		٢٥٠

والخطوة التالية لذلك قراءة الأرقام الاصلية قيمة قيمة ووضع علامة (/) أمام كل فئة تقع بين حدودها تلك القيمة واحدة بعد الأخرى .

وتسهيلاً لعملية العد النهائي يستحسن وضع العلامات في مجموعات خماسية (𝍸) كما هو واضح ولا يتبقى بعد ذلك الا حصر تلـــك العلامات وتسجيل عددها أمام كل فئة . ويسمى الجدول الذي استخدم في توزيع القيم في فئاتها جدول تفريغ (Tallying Table) ، أمـــا الجدول الذي يوضح الفئات والتكرارات دون العلامات المساعدة يسمى جدول تكراري للاختصار . ويمكن تطبيق فكرة التوزيعات التكرارية على البيانات غـير الرقمية اسوة بالرقمية منها. فواضح أن الفكرة الأساسية في أي توزيع تكراري هي تحديد عدد مرات ظهور القيم المختلفة للظاهرة موضوع الدراسة .

فلو فرض أن كان لدينا كشفاً بالحالة الزواجية للعمال السابقين فيمكن تلخيص تلك البيانات في شكل توزيع تكراري غير رقمي بكتابة الحالات

الزواجية المختلفة ثم تحديد عدد العمال الذين يشتركون في كل حالة معينة كما فعلنا في مسألة الأجور بالضبط. ويكون التوزيع التكراري المطلوب هكذا:

الحالة الزوجيـــة	التكرار (عدد العمال)
أعـــزب	٤٨
مـتزوج	١٣١
مطلـق	٤١
أرمـل	٣٠

ويمكن وضع التوزيع التكراري المطلق في شكل توزيع نسبي وذلـك بحساب نسبة تكرار كل فئة من مجموع التكرارات .

التوزيع التكراري المستمر وغير المستمر – المتصل وغير المتصل :

ويجدر بنا أن نشير الى وجودنوعين أساسيين من المتغيرات (Variables) فهناك متغيرات يمكن بطبيعتها أن تأخذ جميع القيم بين حدي مدى التغيـر ويسمى هـــذا النوع « متغيرات مستمرة » (Continuous Variables) ، ومثله الأجور في المثال الذي أشرنا اليه اذ ان الأجر يمكن – نظرياً على الأقل – أن يأخذ جميع القيم بين ٤ ليرات، ٥٢ ليرةوهو مدىالتغير في المثال السابق – فيمكن أن نتصور وجود عامل أجره ٤ ١/٧ ليرة وثالث أجره ٤ ٣/٤ ليرة ورابع أجره ٥ ليرة وخامس أجره ٥ ١/٣ ليرة وهكذا حـــتى نهاية المدى .

وبالمثل اذا كان لدينا توزيعاً تكرارياً لأوزان ألف طالب فان المتغير وهو وزن الطالب متغير مستمر حيث أن وزن كل طالب منهم يمكن ولو نظرياً على الأقل أن يأخذ أي قيمة بين حدي مدى التغير .

وكذلك اذا كان لدينا توزيعاً تكرارياً للأسعار الشهرية للقطن خـــلال العشرين سنة الماضية فان المتغير هنا وهو السعر متغير مستمر اذ لو تبين لنا

أن السعر تغير في المدى بين عشرين ريال ، ٢٠٠ ريال مثلاً يمكن تصور حدوث أي سعر بين هذين الحدين خلال المدة موضوع الدراسة .

وعلى العكس من ذلك يوجد نوع آخر من المتغيرات لا يمكن بطبيعته أن يأخذ جميع القيم التي تقع في حدود مدى التغير وهذا النوع يُطلق عليه اسم المتغيرات غير المستمرة أو الوثابة (Discontinuous Variables) .

فلو فرض ان كان لدينا توزيعا تكرارياً لعدد أفراد اسر حي معين من أحياء بيروت ، فالمتقير هنا هو عدد أفراد الأسر متغير وثاب ، حيث أن عدد أفراد الأسرة لا يمكن أن يأخذ جميع القيم بين حدي المدى . فلو فرضنا أن عدد الأفراد أسر هذا الحي يتروح بين ١ ، ٨ أفراد مثلا ، فيمكن تصور وجود عائلات عدد أفرادها ٢ ، ٣ ، ٤ وهكذا . ولكن لا يمكن تصور وجود عائلات عدد أفرادها ١/٢ ٤ أو ١/٢ ٥ .

ولو كان لدينا توزيعا تكراريا لعدد ثمار التفاح الذي تحمله اشجار صنف معين منها وثبت لنا أن مدى التغير يتروح بين ٥٠ ثمرة ، ٢٠٠ ثمرة مثلا فالمتغير هنا هو عدد الثمار متغير وثاب اذا لا يمكن تصور وجود شجرة تحمل ١/٢ ٧٠ ثمرة أو ١/٤ ٩٣ ثمرة مثلا

ومن البديهي أن جميع التوزيعات التكرارية للقيم غير الرقمية لا يمكن أن يكون المتغير فيها مستمرا ، ففي مثال تقسيم العمال حسب الحالة الزوجية كان المتغير وثابا بطبيعته ولا يمكن أن يكون غير ذلك .

ونلاحظ ان توزيع الطلبة تبعاً للاعمار وان كان يمكن ان يكون توزيعا متصلاً ، الا انه حيث أن الاعمار تكون بالسنوات الكاملة لذلك تكتب فئات العمر على أساس غير متصل .

التوزيع التكراري المفتوح والمغلق :

والتوزيعات التكرارية نوعان الأول مفتوح وهو ما لم يتعين حده الاعلى

أو حده الأدنى أو كليهما لفئتيه المتطرفتين. والثاني مغلق وهو ما عرف حدا فئتيه العليا والسفلى ومثل ذلك التوريع السابق لأحور العمال حيث أن الحد الأدنى للفئةالاولى والحد الأعلىللفئة الأخيرة معروفان وهناك حالات تتطلب طبيعة البيانات فيها عدم تعيين الحد الأدنى للفئة الأولى منها أو الحد الأعلى للفئه الأخيرة اما لاعتبارات عمليـــة حيث تقتضي الاشارة إلى كثير من المفردات غير المتجانسة بأنها تقع في فئة غير محدودة الأطراف ، واما لجهل الباحث بحدود الفئات فمثلا قد لا يستطيـع ممثلوا وزارة الزراعة معرفة المساحات التفصيلية التي يملكهـــا صغار الملاك في قرية معينة على الرغم من معرفة عددهم ، ولذا يطلقون على الفئة الاولى أقل من نصف فدان مثلا جهلا بحقيقة الحد الأدنى للفئة .

ومن الأفضل أن يحاول الباحث جعل التوزيع التكراري الذي يقوم بعمله مقفول الطرفين كلما أمكن ذلك ، اذ ان جعله مفتوحا يترتب عليـه عدم امكانية استخدام بعض المتوسطات – كما سنشير إلى ذلك عنـد الكلام عن المتوسطات – وكذلكيصعب وضع التوزيـع التكراري المفتوح في شكل بياني كامل كما سيظهر عند معالجة هذا الموضوع فيما بعد .

التوزيـع المنتظم وغير المنتظم :

إذا كان مدى الفئات في التوزيـع التكراري متساوي في جميع الفئات فإن التوزيـع يكون منتظماً . ولكن يبدو أحياناً ان القيم تتركز في جزء من التوزيـع بينما تكون مبعثرة في الاجزاء الأخرى ، ولذلك إذا جعلنا التوزيـع منتظما فان كثير من الفنات تكون خالية من التكرارات ولهـذا يستحسن أن يكون التوزيـع غير منتظم بجعل الفئات غير متساوية وذلك بان تكون الفئات في الاجزاء التي تتركز فيها القيم ذات مدى قصير بينما تكون ذات مدى واسع في الأجزاء التي تتبعثر فيها القيم .

تحديد عدد الفئات :

من أصعب الأمور اختيار عدد الفئات التي يجب توزيـــع القيم الأصلية

عدد فئات التوزيع التكراري = ١ + ٣٫٣ × لوغاريتم عدد القيم

عليها اذ ان ذلك يتوقف على اعتبارات عدة . هذا وان كانت الخبرة والمران خير مرشدين لأنسب الوسائل الا انه لا مانــع من ذكر بعض الاعتبارات الأساسية التي لا يجب ان يغفلها الباحث في هذا الشأن .

فلا يجب ان نغالي في الأكثار من عدد الفئات اذ ان ذلك يتنافى والسبب الرئيسي الذي من اجله قمنا بعمل التوزيع التكراري نفسه الا وهو الاختصار والتركيز . ومن الناحية الأخرى لا يجب ان نغالي في الاختصار نفسه ونضع التوزيع في فئتين أو ثلاث اذ ان في ذلك اختصار مخل .

هذا وان كان من الصعب اعطاء قاعدة لهذا الموضوع الا انه لا يجوز على العموم ان يقل عدد فئات التوزيع عن عشرة ولا يزيد عن عشرين فئة تبعا لكثرة عدد القيم المراد تلخيصا أو قلتها .

ومما لا بأس به الأشارة هنا إلى قاعدة رياضية لتحديد عـــدد فئات التوزيعات التكرارية ذات القيم متوسطة العـــدد أي بين مائة وألف قيمة ويطلق على هذه القاعدة ، قاعدة ستيرجس Sturges Rule وهي :

عدد فئات التوزيع التكراري = ١ + ٣و٣ × لوغارثم عدد القيم .

وعند تقسيم القيم الى فئـــات يجب ان نلتفت إلى قاعدتين أساسيتين ، الأولى هي الا تكون الفئات متباعدة بمعنى ان يكون هناك هوّة أو مسافة بين الحد الأعلى للفئة والحد الأدنى للفئة التالية لها . فلو كنا في صدد تقسيم عدد من العمال حسب ساعات العمل التي يشتغلها كل واحد منهم اسبوعيا فلا يجب وفقا لتلك القاعدة – ان يكون التقسيم كالآتي :

٤٠ ساعة إلى ٤٥ ساعة
٤٦ ساعة إلى ٥٠ ساعة
٥١ ساعة إلى ٥٥ ساعة

إذا أن هناك مسافة بين الحد الأعلى للفئة الأولى والحد الأدنى للفئة الثانية وهكذا ، فيقال أن هناك تباعد بين تلك الفئات – فاذا اتبعنا هذه الطريقة

في تقسيم الفئات يكون من الصعب تحديد الفئة التي ينتمي اليها عامل عدد ساعات عمله الاسبوعية ٤٥ ساعة ونصف أو آخر عدد ساعات عمله ٥٠ ساعة وربع. هذا ولا يجب أن نغفل عند تطبيق هذه القاعدة طبيعة المتغير نفسه ودرجة التقريب فيه. فلو كان المتغير وثّاباً مثلاً فمما لا غبار عليه أن نجعل الفئات متباعدة إذ انه من الخطأ أن نجعل الفئات غير ذلك.

فلو كنا بصدد تقسيم عدد كبير من الشقق بحسب عدد الغرف التي تحتوي عليها كل منها فلا مانع بل لا بد من جعل الفئات متباعدة هكذا:

غرفتان إلى أربع غرف
خمسة غرف إلى سبعة غرف
ثمانية غرف إلى عشرة غرف

فقد يظن القارىء لاول وهلة أن هناك تباعد بين تلك الفئات خلافاً للقاعدة ولكنه لو تذكر أن المتغير هنا وثاب وانه لا يمكن أن يأخذ قيما تقع بين أي من تلك الفئات المذكورة لتحقق من صحة التقسيم.

وبالمثل قد يكون السبب في التباعد بين الفئات التقريب الحسابي في قيم المتغير الأصلي فلو فرض اننا نقوم بعمل توريع تكراري لطول محيط صدر مجموعة من الطلبة بالسنتمتر مثلا وان محيط الصدر كان يقاس بأقرب سنتمتر فلا مانع اذا من وضع الفئات هكذا:

١٢١ سنتمتر إلى ١٣٠ سنتمتر
١٣١ سنتمتر إلى ١٤٠ سنتمتر
١٤١ سنتمتر إلى ١٥٠ سنتمتر

فلا يمكن قبول الانتقاد الذي سبق أن وجهناه للفئات المتباعدة في هذه الحالة إذ أن من كان محيط صدره ١٣٠ سنتمتر يقع في الفئة الأولى ومن كان محيط صدره ١٣٠ سنتمتر وكسر لا بد أن يسجل طول محيط صدره لأقرب

سنتمتر فاما ان يكتب ١٣٠ سنتمتر ويقع لذلك في الفئة الأولى أو ١٣١ سنتمتر ويقع لذلك في الفئة الثانية .

أما إذا كان المتغير مستمراً فيكون من الخطأ ان نجعل هناك أي تباعد بين الفئات بمعنى ان الحد الأدنى لكل فئـــة هو نفسه الحد الأعلى للفئة السابقة .

والقاعدة الثانية التي يجب أن نتبعها عند تقسيم القيم إلى فئات هي الا تكون الفئات متداخلة Overlapping Groups بمعنى أن يكون جزء من فئة داخل في فئة أخرى سابقة أو لاحقة . فلو كنا بصدد تقسيم الشركات المساهمة حسب معدلات للربح التي توزعها كل شركة على حاملي أسهمها وقسمنا التوزيع إلى الفئات الآتية :

٢ ٪	إلى	٥ ٪
٥ ٪	»	٨ ٪
٨ ٪	»	١١ ٪

فيحتار الباحث في وضع الشركة التي وزعت ربحاً على مساهميها بنسبة ٥٪ بين الفئتين الأولى والثانية . وبالمثل من الخطأ اعتبار الشركة التي وزعت ربحاً بمعدل ٨٪ من بين وحدات الفئة الثانية دون الثالثة اذ ان هاتين الفئتين تتنازعاها وذلك بسبب تداخل الفئات بعضها في البعض الأخر .

طرق كتابة الفئات :

نظراً للاعتبارات السابقة يجب كتابة الفئات باحدى الطرق الآتية :

٢ ٪	وأقل	من	٥ ٪	
٥ ٪	»	»	٨ ٪	
٨ ٪	»	»	١١ ٪	وهكذا

واختصاراً للكتابة جرت عادة الاحصائيين على كتابة هذه الفئات نفسها هكذا :

٢ ٪ –   وتقرأ ٢ ٪ وأقل من ٥ ٪
٥ ٪ –   »   ٥ ٪ »   »   ٨ ٪
٨ ٪ – ١١ ٪ »   ٨ ٪ »   »   ١١ ٪

ويمكن كتابة الفئات بصورة أخرى هكذا :

أكثر من ٢ ٪ إلى ٥ ٪ وتكتب باختصار – ٥٪
أكثر من ٥ ٪ إلى ٨ ٪ وتكتب باختصار – ٨٪
أكثر من ٨ ٪ إلى ١١ ٪ وتكتب باختصار – ١١٪

هذا ولو ان تلك الطريقة صحيحة إلا أنها ليت شائعة الاستعمال كالطريقة الأولى .

في كل من الطريقتين مدى الفئة ٣ ومراكز الفئات ٣,٥ ، ٦,٥ ، ٩,٥ . والفرق بينهما ان الحد الأعلى للفئـــة الأولى مثلاً هو أقل بقليل من ٥ ٪ في الطريقة الأولى و ٥ ٪ تماماً في الطريقة الثانية – كذلك الحد الأدنى للفئة الأولى هو ٢٪ أي ان الرقم ٢٪ داخل فيالفئة الأولى تبعاً للطريقة الأولى بينما في الطريقة الثانية الحد الأدنى للفئـــة يزيد قليلاً عن ٢٪ والصورة الثالثة لكتابة الفئات هي هكذا :

٣,٥

٦,٥

٩,٥

وهي تعني الفئات التي مراكزها ٣,٥ ، ٦,٥ ، ٩,٥ . ويمكن معرفة المدى بحساب الفرق بين كل مركز وآخر وهو ٣ في هذه الحالة ، فاذا أردنا أن نحدد الفئات التي لها هذه المراكز نقسم المدى على ٢ أي $\frac{٣}{٢}$ = ١,٥ ،

بعد ذلك تطرح ١,٥ من كل مركز نحصل على الحد الأدنى للفئة ونجمع ١,٥

على كل مركز نحصل على الحد الأعلى للفئة وتبعاً لذلك تكون الفئات التي مراكزها الأرقام السابقة هي ٢ – ٥ ، ٥ – ٨ ، ٨ – ١١ .

ومن أهم ما يجب التنويه بـه ونحن في صدد اختيار فئات التوزيع التكراري العمل على جعل جميع الفئات متساوية المدى ما أمكن وسوف نرى عند الكلام على موضوع المتوسطات ان في تساوي جميع الفئات اختصار كبير في العمل الحسابي كما أن في عدم التساوي حرمان من امكان استخدام بعض أنواع المتوسطات الهامة .

ويجب أن نذكر بتلك المناسبة ما سبق أن أشرنا اليه وهو ان هناك حالات تستدعي طبيعة البيانات جعل الفئات غير متساوية . فمن أراد أن يدرس التوزيع التكراري للنسب المئوية للمتوفين في فئات العمر المختلفة بين سكان بلد معين بحسب العمر عند الوفاة لا يمكن أن يجعل التوزيع متساو في الفئات . فمن المعروف أن نسب المتوفين في الأعمار المتوسطة أي بين عشرين سنة وخمسين سنة تكاد تكون واحدة فلا يتطلب الأمر التفصيل الدقيق في تلك المرحلة من العمر ، ولكنه من المعروف أيضاً ان درجة تعرض الأطفال وخصوصاً الرضع منهم أي الذين لم يبلغوا الواحدة بعد سريعة التذبذب من سنة الى أخرى بل ومن شهر الى آخر ، الأمر الذي يستدعي من الباحث المدقق العمل على جعل فئات التوزيع قصيرة في أول التوزيع وطويلة في وسطه ثم صغيرة في نهايته . أما فيما عدا تلك الحالات وأمثالها يجب أن يحاول الباحث أن يختار الفئات بحيث تكون كلها متساوية . وهناك اعتبار آخر يجب الالتفات اليه عند اختيار فئات التوزيعات التكرارية وهو أن مركز كل فئة سيمثل جميع الأفراد الذين يقعون بين حدودها وهذا ما يتطلب منا يقظة وحذر شديد في اختيار تلك الحدود اذ لا بد من الالتفات الى طبيعة البيانات الأصلية وطريقة تركزها .

فمن أراد أن يكون توزيعاً تكرارياً لأعمار فئة كبيرة من السكان لا بد أن يتذكر تقريب الأرقام عند اعطاء السن ، فمن كان عمره الحقيقي ١٩ سنة

أو ٢١ سنة مثلا سيجعله ٢٠ سنة . أما الذين أعمارهم ٣٧ سنة أو ٣٨ سنة أو ٣٩ سنة سيسجلونه ٤٠ سنة ولذا وجب اختيار الفئات كالآتي :

١٥	وأقل من	٢٥
٢٥	« «	٣٥
٣٥	« «	٤٥ وهكذا .

فواضح ان الذين اعتبروا في الفئة الأولى لأنهم سجلوا أعمارهم ٢٠ سنة هم في الواقع من ١٥ الى ٢٤ سنة فكأن التقسيم الى تلك الفئات صحح بقدر الأمكان الخطأ الغير المقصود في تسجيل العمر .

ومما هو جدير بالذكر أن مراكز الفئات يمكن تحديدها بحساب متوسط الحدين الأول والأعلى لكل فئة ، فمثلاً في التوزيع ٧ — ، ٩ — ، ١١ — ، يكون مركز الفئة الأولى $\frac{٧ + ٩}{٢}$ =٨ ومركز الفئة الثانية $\frac{٩ + ١١}{٢}$ = ١٠ وهكذا . وهذه القاعدة عامة نطبقها مهما اختلف نوع التوزيع النكراري مع ملاحظة ان التوزيع المفتوح لا يمكن حساب مركز فئته المفتوحة لعدم معرفة أحد حديها

التوزيعات التكرارية المتجمعة :

من التوزيع التكراري العادي يمكن معرفة عدد الوحدات التي تقع في داخل كل فئة مثلاً نعرف عدد الاشخاص الذين يترارح طولهم بين ١٦٠ سم ، ١٦٥ سم . ولكن قد يطلب أحياناً معرفة عدد الطلبة الذين يقل طولهم عن ١٦٥ سم أو يزيد طولهم عن ١٦٠ سم ، ولا يمكن الاجابة بسهولة وبشكل مباشر على تلك الأسئلة بمجرد النظر الى التوزيع العادي . ولكي نتمكن من الاجابة على تلك الاسئلة وغيرهـــا لا بد من وضع التوزيع في شكل جديد سنطلق عليه اسم التوزيع التكراري المتجمع .

والفكرة الاساسية في التوزيعـــات التكرارية المتجمعة هي تجميع التكرارات امام الحد الاعلى لكل فئة وفي هذه الحالة يكون التوزيع التكراري متجمع صاعد حيث أن التكرارات في صعود مستمر ، أو تجميعها أمام الحد الادنى لكل فئة ابتداء من أسفل التوزيع وفي هذه الحالة يكون التوزيع متجمع هابط حيث تكون التكرارات في هبوط مستمر .

توزيعان تكراريان متجمعان صاعد وهابط

توزيــع تكراري بسيط		توزيــع تكراري متجمع			
فئات الأجر	التكرار (عدد العمال)	الحدود العليا للفئات	التكرار المتجمع الصاعد	الحدود الدنيا للفئات	التكرار المتجمع الهابط
٤ —	٦	أقل من ٨	٦	٤ وأكثر	٢٥٠
٨ —	١٤	» » ١٢	٢٠	٨ » »	٢٤٤
١٢ —	٣١	» » ١٦	٥١	١٢ » »	٢٣٠
١٦ —	٥٤	» » ٢٠	١٠٥	١٦ » »	١٩٩
٢٠ —	٤٣	» » ٢٤	١٤٨	٢٠ » »	١٤٥
٢٤ —	٣٦	» » ٢٨	١٨٤	٢٤ » »	١٠٢
٢٨ —	٢٧	» » ٣٢	٢١١	٢٨ » »	٦٦
٣٢ —	١٩	» » ٣٦	٢٣٠	٣٢ » »	٣٩
٣٦ —	١٦	» » ٤٠	٢٤٠	٣٦ » »	٢٠
٤٠ —	٦	» » ٤٤	٢٤٦	٤٠ » »	١٠
٤٤ —	٣	» » ٤٨	٢٤٩	٤٤ » »	٤
٤٨ — ٥٢	١	» » ٥٢	٢٥٠	٤٨ » »	١

والفائدة العملية لهذا النوع من التوزيع التكراري اننا نستطيع أن نعرف بمجرد النظر الى الجدول السابق عدد العمال الذين يتقاضون أجور أقل من أجر معين او أعلى من أجر معين ، فمثلا يتضح ان هناك ٥١ عامل يتقاضون اجراً اقل من ١٦ ليرة وان ١٩٩ عامل يتقاضون أجرا يزيد عن ١٦ ليرة وبالطبع يمكن معرفة نسبة هذا العدد الى المجموع ووضعها في خانة مقابلة للجدول .

كذلك من هذا الجدول يمكن معرفة الأجر الذي يتقاضاه عدد معين من العمال ، فمثلا من التوزيع المتجمع السابق يتضح ان الأجر الذي يتقاضاه ٢٣٠ من العمال هو أقل من ٣٦ ليرة ، كما يمكننا أن نعرف ان هذا العدد من العمال يتقاضى أكثر من ١٢ ليرة .

ونلاحظ انه إذا كان لدينا توزيع تكراري متجمع فيمكن بعملية عكسية (عملية طرح) أن نحصل على التوزيع التكراري الأصلي ، فمثلاً :

الحد الأعلى للفئة	التكرار المتجمع الصاعد	الفئات	التكرار
أقل من ٦٤,٥	٣	- ٦٤,٥	٣
« « ٧٤,٥	٢٤	٦٤,٥ —	٢١
« « ٨٤,٥	١٩٢	٧٤,٥ —	٧٨
« « ٩٤,٥	٢٨٤	٨٤,٥ —	١٨٢
« « ١٠٤,٥	٥٨٩	٩٤,٥ —	٣٠٥
« « ١١٤,٥	٧٩٨	١٠٤,٥ —	٢٠٩
« « ١٢٤,٥	٨٧٩	١١٤,٥ —	٨١
« « ١٣٤,٥	٩٠٠	١٢٤,٥ —	٢١
« « ١٤٤,٥	٩٠٥	١٣٤,٥ — ١٤٤,٥	٥

كذلك نلاحظ أن التوزيع المتجمع المطلق سواء كان صاعداً أم هابطاً يمكن تحويله إلى توزيع متجمع نسبي وذلك بحساب نسبة كل تكرار متجمع من مجموع التكرارات الذي يكون هو آخر تكرار متجمع .

التوزيع التكراري المزدوج :

يذكر القارىء ان من أهم فوائد التوزيع التكراري تلخيص البيانات في شكل يسمح للعقل بالوقوف على أهم خصائص القيم التي نريد اختصارها . ولكن حتى الآن كنا نقتصر على معالجة عدد من قيم متغير واحد .

ولكن لو كان لدينا بيانات تخص ظاهرتين بينهما علاقة مثلاً بيانات عن عمر الزوج وعدد الأولاد الذين عنده ، أو عن الأسعار المختلفة لسلعة ما والكمية التي تعرض بهذه الأسعار ، أو عن علامات بعض الطلبة في اختبار الذكاء وعلاماتهم في اختبار علم من العلوم ، أو عن كمية المطر المتساقط في شهر يوليو ومتوسط كمية الانتاج من القمح ، أو عن أطوال مجموعة من الطلبة وأوزانهم — في مثل هذه الحالات لا يمكن اختصار البيانات بواسطة التوزيع التكراري العادي حيث يكون لدينا نوعان من البيانات أمام كل وحدة ، وقد تكون هذه البيانات مقاسة بوحدات مختلفة . فاذا حاولنا اختصار كل نوع منها في توزيع خاص بها سوف يكون لدينا توزيعان منفصلان عن بعضهما الأمر الذي لا يساعدنا في دراسة العلاقة بين الظاهرتين . ولهذا نحاول تبويبها في جدول واحد ، هو جدول التوزيع التكراري المزدوج ، فمثلاً إذا كان لدينا البيانات الآتية عن أجور بعض من العمال وعدذ أولاد كل منهم فانه من الممكن اختصار تلك البيانات في جدول توزيع تكراري مزدوج وذلك بتحديد الفئات (الأرقام) الخاصة بالعمر والفئات الخاصة بعدد الاطفال ثم نبدأ بتفريغ البيانات السابقة في الفئات المزدوجة التي تقابل بعضها ويكون التفريغ بوضع علامات في كل خانة تنم عن الفئة المزدوجة التي ينحصر بين مداها كل زوج من القيم . بعد ذلك نلخص جدول التفريغ في جدول تكراري مزدوج بترجمة العلامات إلى أرقام كما فعلنا في حالة التوزيع التكراري البسيط تماماً .

١٦-٠، ٢٠-٣، ٨-١، ١٢-٠، ١٣-٠، ٢٩-٠، ٤٣-٥، ٣٤-٠، ٢٣-٢، ٢٠-٢ ،
٢٤-١، ٤٥-٣، ٣٦-٤، ٢٩-١، ٣٥-٦، ٢٥-٠، ٢٠-١، ١٤-٠، ١١-٠، ١٠-٠ ،
١٦-٠، ١٢-٣، ٤٠-٢، ٣٢-٥، ٢٤-٠، ٢٠-١، ١٨-٢، ١٥-١، ٨-٠، ٢٢-٠ ،
١٨-٠، ٢١-١، ٢٦-٠، ٣١-٢، ٣١-٤، ١٨-٦، ٢١-٣، ٢٦-٣، ١٥-٢، ٢٧-٢ ،
٢٠-٢، ١٥-٣، ١٧-٠، ٢١-٥، ١٥-٢، ١٧-٠، ١٩-٦، ٢٩-٣، ٢٥-٢، ٢٣-٧ ،

٢٠-١، ١٩-٥، ١٧-٢، ١٤-٠، ١٨-٣، ٢٥-٣، ٣٠-٣، ٣٥-٤، ٢٣-١، ١٥-١ ،
١٥-٠، ٢١-٠، ١٠-٠، ٢٥-٢، ١٣-١، ٣٠-٢، ١٨-٢، ٣٨-٥، ٢٠-٤، ٢٤-٢ ،
٢٥-٣، ٢٩-٣، ٣٤-١، ١٧-١، ٢٢-١، ٢٥-٥، ٢٧-٧، ٢٨-٥، ٣٥-٥، ٣٠-٣ ،
٢٥-٤، ١٨-٠، ٢٩-٤، ١٦-٢، ٣٨-٦، ١٠-٠، ١٦-٠، ٢٠-٠، ١٣-٠، ٢٣-١ ،
٣٦-٣، ٣٢-١، ٢٠-٠، ١٦-٠، ٢٠-٥، ١٧-٠، ٥-٠، ٢٧-٢، ٢٢-٢، ١٨-٠ ،

٣٩-١، ١٩-٤، ١٠-٠، ٧-٠، ٤-٠، ٣٧-٣، ٥٠-٤، ١٦-٠، ٣٣-٥، ٣٥-٧ ،
٢٧-١، ١٦-٠، ٩-١، ٦-٠، ٤٥-٦، ٣٩-٧، ١٥-١، ١٧-٠، ١٧-٠، ٣٠-٦ ،
٢٧-١، ٣٦-٢، ٤٥-٦، ١٢-٠، ١٦-٠، ٨-٠، ٢٠-٦، ٢٢-١، ٢٤-٢، ٢٨-٥ ،
٣٣-٦، ٣٧-٢، ٤٢-٦، ٢٢-٠، ١٨-٠، ١٥-٠، ٢٧-٢، ٢٢-٢، ١٨-٠، ١٥-١ ،
١١-٠، ١٧-٢، ١٩-٣، ١٤-٠، ٢١-١، ٣٠-٥، ٢٥-٣، ٢١-٢، ١٩-٠، ١٦-٠ ،

١١-٠، ٢٠-١، ١٩-٠، ٢٥-٦، ٢١-١، ٣٠-١، ٢٦-٠، ٢٠-٠، ١٥-٢، ١٩-٠ ،
١٤-١، ٢٥-٠، ١٣-٠، ١٦-٠، ١٧-٠، ١٩-٢، ٢١-٠، ٢٣-١، ١٥-٠، ٢٥-٠ ،
١٧-٣، ٣٥-٣، ٢١-٢، ١٠-١، ١٩-٠، ٢٥-٢، ١٥-٢، ١٧-٠، ١٧-٠، ٢٠-٣ ،
١٩-٢، ٢٩-٢، ٣٥-٤، ٢٧-٠، ٢٤-٣، ٩-١، ١٧-٠، ١٩-٠، ٢٥-٠، ٣٥-٢ ،
٣٣-٩، ٢٨-٧، ٢٠-٠، ١٢-١، ٨-٠، ٥-٠، ١٥-١، ٢١-٠، ٣٥-٤، ٣٥-١ ،

٣٠-٠، ٣٥-٣، ٢٥-١، ٢٦-١، ٢٠-٢، ١٩-٢، ٣١-٣، ١٧-٠، ١٤-٢، ١٥-١ ،
٢٢-٧، ٢٢-٢، ٢٢-١، ١٧-٠، ١٢-٢، ٢٠-٣، ٣٠-٥، ١٩-٠، ٢١-٢، ٥-٠ ،
١٩-٠، ١٩-٠، ٢٠-٠، ٢٣-٢، ١٦-٠، ١٤-٠، ١٠-٠، ٣٠-٤، ٢٥-٢، ٢١-٢ ،
١٧-٠، ٢٠-٧، ١٩-٠، ٢٣-٠، ٢٥-٨، ١٨-٠، ٤٠-٤، ٤٠-١، ٤١-٥، ٢٤-٤ ،
١٥-٠، ٢٢-٠، ٣٤-٣، ٢٨-٣، ٢٧-٠، ١٥-٠، ٢٨-٣، ٣٠-٨، ١٥-١، ١٩-٠ ،

عدد الأطفال	٤–	٨–	١٢–	١٦–	٢٠–	٢٤–	٢٨–	٣٢	٣٦	٤٠	٤٤	٤٨–	المجموع
	فئــات الأجــور (بالليرة)												
صفر	卌ǀ	卌 卌	卌 卌 卌	卌 卌 卌 ǀǀǀǀ	卌 卌 ǀǀ	卌 ǀǀǀǀ	ǀǀǀ	ǀǀ	–	ǀ	–	–	٧٧
١	–	ǀǀǀǀ	卌 卌	卌 卌 卌 卌 ǀ	卌 卌 ǀǀ	卌 ǀ	ǀǀ	ǀǀǀ	–	–	–	–	٥٨
٢	–	–	卌	卌 ǀǀǀ	卌 卌 ǀ	卌 ǀǀǀǀ	ǀǀǀ	ǀ	ǀǀ	ǀ	–	–	٤٠
٣	–	–	ǀ	ǀǀǀ	ǀǀǀǀ	卌	卌 ǀǀ	ǀǀǀǀ	ǀǀ	–	ǀ	–	٢٧
٤	–	–	–	ǀ	–	ǀǀ	卌	ǀǀǀ	ǀǀ	ǀ	–	ǀ	١٥
٥ فأكثر	–	–	–	ǀǀ	ǀǀǀǀ	卌	卌 ǀǀ	卌 ǀ	ǀǀǀǀ	ǀǀǀ	ǀǀ	–	٣٣
المجموع	٦	١٤	٣١	٥٤	٤٣	٣٦	٢٧	١٩	١٠	٦	٣	١	٢٥٠

عدد الأطفال	٤–	٨–	١٢–	١٦–	٢٠–	٢٤–	٢٨–	٣٢–	٣٦–	٤٠–	٤٤–	٤٨–	المجموع
	فئــات الأجــور (بالليرة)												
صفر	٦	١٠	١٥	١٩	١٢	٩	٣	٢	–	١	–	–	٧٧
١	–	٤	١٠	٢١	١٢	٦	٢	٣	–	–	–	–	٥٨
٢	–	–	٥	٨	١١	٩	٣	١	٢	١	–	–	٤٠
٣	–	–	١	٣	٤	٥	٧	٤	٢	–	١	–	٢٧
٤	–	–	–	١	–	٢	٥	٣	٢	١	–	١	١٥
٥ فأكثر	–	–	–	٢	٤	٥	٧	٦	٤	٣	٢	–	٣٣
المجموع	٦	١٤	٣١	٥٤	٤٣	٣٦	٢٧	١٩	١٠	٦	٣	١	٢٥٠

وفضلا عن تلخيص البيانات الأصلية ووضعهـــا في شكل يسمح بتفهم العلاقات فيما بينها فان الجدول السابق يمكننا من الوقوف على طبيعة كل من المتغيرين على حده فيمكن مثلاً معرفة التوزيع التكراري للأجور بصرف النظر عن عدد الأطفال، وكذلك يمكن الوقوف على طبيعةالتوزيع التكراري لعدد الأطفال بصرف النظر عن توزيع الأجور .

ويمكن في الواقع استخلاص عدة توزيعات تكرارية بسيطة من واقع هذا التوزيع المزدوج ، فمثلا يمكن معرفة التوزيع التكراري لأطفال العمال الذين يتقاضون من ٢٤ قرشاً الى أقل من ٢٨ قرشاً وذلك باستخدام تكرارات العمود الأول والسابع من الجدول المزدوج ، وبالمثل يمكن استنباط التوزيع التكراري لأجـــور العمال الذين لديهم خمسة أطفال فأكثر وذلك باستخدام السطرين الأول والسابع من نفس الجدول .

وفضلاً عن ذلك فان جدول التوزيع المزدوج يمكننا من الوقوف على حقيقة العلاقة بين المتغيرين ان كانت هناك ثمة علاقة ، فلو ألقينا نظرة سريعة على تكرارات خانات الجدول المزدوج لأمكننا القول بأنه كلما زاد الأجر للعامل زاد تبعاً لذلك (على العموم) عدد أطفال هذا العامل . فنرى مثلاً أن هناك ستة عمال يتقاضون اجراً من أربعة قروش إلى أقل من ثمانية قروش وليس لديهم أطفالا وأن هناك أربعة عمال أجرهم من ثمانية قروش وأقل من ١٢ قرشاً ولديهم طفل واحد . وهذا يفسر حقيقة كون الجدول خالياً من التكرارات عندما يكون عدد الأطفال كبيراً والأجور قليلة وكذلك عندما تكون الأجور كبيرة وعدد الاطفال قليلا ، وبعبارة اخرى يمكن القول بان هناك علاقة أو ارتباط بين الظاهرتين – الأجر وعدد الأطفال – بمشاهدة انتظام تكرارات الجدول في شكل شريط مائل من الطرف الأيمن في أعلى

الجدول إلى الطرف الأيسر في أسفله ان كانت العلاقة مضطردة أي الارتباط موجباً ، وينعكس وضع هذا الشريط ان كانت العلاقة عكسيةأي الارتباط سالبا (هذا على فرض ان التوزيعين وضعا في الجدول بشكل تصاعدي) ، ولهذا السبب يطلق الاحصائيون على هذا الجدول اسم « جدول ارتباط » .

وانتظام التكرارات في هذا الشريط دليل على وجود الارتباط ، أما تحديد قيمة الارتباط نفسه فسنرجىء بحثه عندما ندرس بالتفصيل موضوع الارتباط .

الفصل الرابع

التوضيح البياني

رأينا كيف يستخدم التبويب في وضع عدد كبير من البيانات في صورة منظمة واضحة تساعد على استيعاب الموضوع الذي نقوم بدراسته . على ان هذه الصورة قد لا تكفي احيانا للتوضيح خاصة وان بعض الناس يجدون صعوبة كبيرة في ادراك مدلولات الارقام التي تعرض عليهم في جداول ، ولهذا يحاول الاحصائي توضيح البيانات بطريقة اخرى وهي الرسوم البيانية، حيث ان كثيرا من هذه الرسوم يساعدنا احيانا في تكوين فكرة سريعة ودقيقة عن كثير من البيانات المعقدة . هذا فضلا عن ان بعض الرسوم تساعدنا في اجراء التحليل الاحصائي .

وتختلف وسائل العرض البياني باختلاف البيانات التي لدينا ، وذلك لأن الرسم البياني يجب ان يصمم بحيث يبرز الفكرة الاساسية التي يرغب مصمم الرسم في ابرازها ، ولهذا تكون اول خطوة في التوضيح البياني هي اختيار الرسم الذي يناسب بيانات معينة حتى تبرز الفكرة التي تدور حولها هذه البيانات . ونلاحظ ان طرق العرض البياني قد تطورت وتشعبت بحيث اصبحت فنا له قواعده واصوله . ولا يتسع المجال في هذا الكتاب لاستعراض الطرق المختلفة التي يمكن ان تستخدم في تجميل الرسوم البيانية حتى تزيد جاذبيتها . وبذلك سنقتصر على دراسة الطرق الرئيسية على ان يؤخذ في

الاعتبار دائماً ان الرسم البياني هو فن قبل ان يكون عملا احصائيا ولهذا لا يجب ان نبخل على الرسم بكل ما يساعدنا على زيادة جاذبيته ووضوحه .

الرموز البيانية :

ولعل أول طريقة للتوضيح البياني هي الرموز البيانية وهي طريقة تستخدم غالبا في النشرات التي تنشرها الحكومة والهيئات الخاصة بقصد تفهيم افراد الجمهور بعض الحقائق الرقمية ، وقد ازداد استعمال هذه الطريقة في بعض الدول التي يكون فيها حقاً للجمهور أن يتعرف على الحقائق المختلفة التي تمس حياته الاقتصادية والاجتماعية . وتبعاً لهذه الطريقة تمثل الظاهرة موضوع التوضيح بصورة ترمز اليها (صورة مصنع للدلالة على المؤسسات الصناعية مثلا) ثم بتغيير حجم هذه الصورة أو تغيير عدد وحداتها يمكن التعبير عن الاتجاه الرقمي لهذه الظاهرة . الا ان هذه الطريقة ينقصها الدقة في التوضيح خاصة إذا تذكرنا أن الصورة مهما كانت دقيقة لا يمكن أن تعبر بدقة عن التغير في الأرقام ، ولهذا فهي تناسب الاغراض التي تستخدم فيها فقط حيث يكون الهدف من استعمالها هو اعطاء فكرة عامة بقصد الدعاية عن الاتجاه الرقمي لظاهرة ما . أما بالنسبة للأغراض الاخرى التي تكون للدقة فيها أهمية كبيرة يحاول الاحصائي استخدام بعض الطرق الاخرى في التوضيح

شكل الأعمدة :

يستعمل شكل الاعمدة في توضيح قيم ظاهرة ما في عدة فترات زمنية وذلك لابراز التغير الذي حدث فيها ، وكذلك في توضيح قيم الأوجه المختلفة لظاهرة معينة لابراز المقارنة بين هذه الأوجه . على اننا نلاحظ انه عند استعمال هذه الطريقة في التوضيح البياني يجب أن لا نغالي في استعمالها حيث أن كثرة الاعمدة في الرسم تجعله يبدو مزدحماً غير مقبولا .

ويتكون شكل الاعمدة من مستطيلات ذات سمك واحد ومنفصلة عن بعضها بحيث تتناسب ارتفاعاتها مع قيم الظاهرة في الفترات الزمنية المختلفة أو مع قيم الاوجه المختلفة للظاهوة . وعندما يرسم هذا الشكل على أساس مستطيلات افقية وليست راسية تكون أيضاً المستطيلات ذات سمك واحد ومنفصلة عن بعضها بحيث تتناسب اطوالها مع القيم المختلفة للظاهرة . واذا كانت بعض البيانات التي تمثل الظاهرة موجبة والبعض سالبة ترسم القيم الموجبة في الاتجاه الموجب من الرسم البياني والقيم السالبة في الاتجاه السالب بحيث يكون المحور الصفر فاصلا بينها ، وذلك مثل توضيح التغير في الميزان التجاري أو في ميزان المدفوعات أو في ميزانية احدى المؤسسات حيث تكون القيم لهـــذه الظواهر في بعض السنوات موجبة وفي بعض السنوات سالبة .

ومن الواضح انه يحسن ابراز هذه المستطيلات باللون الاسود أو بتظليلها بواسطة خطوط متوازية مرسومة داخلها ، وعندما تمثل المستطيلات فيما موجبة واخرى سالبة يحسن تلوينها بالوان مختلفة أو تظليلهـــا بطريقتين مختــلفتين .

وفي هذا النوع من الرسومات البيانية يجب أن يبـــدأ المقياس المدرج من الصفر لأننا إذا بدأنا من أي قيمة أخرى فسوف تختلف النسب بين ارتفاعات الأعمدة أو أطوال المستطيلات عن النسب الحقيقية للأرقام التي تمثلهـــا هذه المستطيلات . ولتوضيح ذلك نفترض ان القيم التي نريد إظهارها في الرسم هي ٣٠٠ ، ٢٠٠ ، ١٠٠ ، أي النسبة الحقيقية بينها هي ٣ : ٢ : ١ فإذا بدأنا بالرقم ٥٠ وليس بالرقم صفر على المقيـــاس المدرج وجعلنا كل سم = ٥٠ وحدة — فإن ارتفاع العمود الأول يكون ٥ سم وارتفاع العمود الثاني ٣ سم وارتفاع العمود الثالث ١ سم أي تكون النسبة بين ارتفاعات المستطيلات هي : ٥ : ٣ : ١ وهي تختلف اختلافاً كبيراً عن النسبة بـين الأرقام التي

توضحها هذه المستطيلات . كذلك يحسن ان تكون المسافة بين كل مستطيل وآخر أو بين كل عمود وآخر نصف المقياس الذي يـدل على سمك العمود أو المستطيل . ومن الأفضل عدم كتابة الارقام التي تمثلها الأعمدة فوق الأعمدة حيث ان ذلك يبالغ في طول الأعمدة بالإضافة إلى جعل الرسم يبدو منفراً .

ويمكن أن تكون الظاهرة ذات قيمتين أو أكثر في كل فترة من الفترات الزمنية ، وفي هذه الحالة إذا كان من المنطق جمع هاتين القيمتين بحيث يكون المجموع دالاً على القيمة الكلية للظاهرة يرسم عموداً لكـــل فترة ويقسم إلى الجزئين بحيث يكون الارتفاع الكلي للعمود معبراً عن القيمة الكلية للظاهرة وارتفاع كل جزء منه معبراً ومتناسباً مع قيمة كل جزء . ويمكن استعمال هذه الطريقة مهما كان عدد الأجزاء التي تتكون منها الظاهرة وسواء أردنا التوضح باستعمال الأعمدة الرأسية أو المستطيلات الأفقية .

أما إذا لم يكن من المنطق جمع هذه الأجزاء يرسم لكل فترة عمودين متلاصقين ومتساويين في القاعدة ومختلفين في الارتفاع تبعاً لاختلاف القيم التي تمثلها هذه الأعمدة . فإذا أردنا توضيح الصادرات والواردات في عدة فترات زمنية لا يكون من المنطق جمعها في عمود واحد ، بل المنطق هو في توضيحها بعمودين متلاصقين حتى يبرز الفرق وهو الذي يدل على الميزان التجاري. ويمكن استعمال هذه الطريقة مهما كان عدد الظواهر التي نريد إظهارها في كل فترة من الفترات . وسواء أردنا التوضيح باستعمال الاعمدة الرأسية أو المستطيلات الأفقية .

ومن الواضح أنه في هاتين الحالتين لا يجب أن نكثر من الأعمدة في الرسم حتى يبدو مكتظاً بها فيضيع بذلك الهدف الأساسي من التوضيح . وبشكل عام لا يجب أن نعرض أكثر من ثلاث ظواهر عن طريق رسم أعمـــدة أو مستطيلات متلاصقة حيث انه كلما زاد عدد الأنواع المختلفة من الأعمدة كلـما أصبح الرسم معقداً وبذلك تقل فائدته .

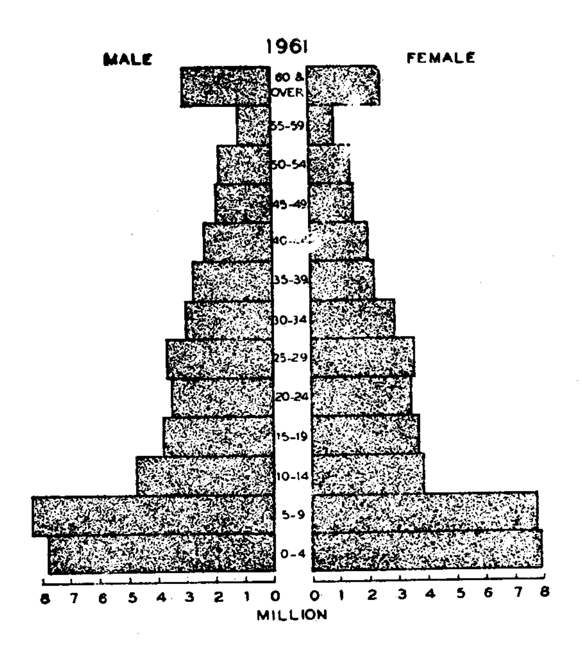
كذلك يجب التميز بين أجزاء الأعمدة أو بين الأعمدة التي تمثل الظواهر المختلفة في كل فترة بالألوان أو التظليل ، على أنه لا يجب أن نلجأ الى الألوان إلا اذا كان لدينا الخبرة في استعمالها والا فان الرسم سوف يبدو قبيحاً منفراً.

وكذلك يمكن أن يستعمل شكل الاعمدة في توضيح التغير الذي حدث في عدة ظواهر بين فترتين زمنيتين (عامين مثلاً) ، ويكون ذلك بجعل المحور الرأسي للرسم في الوسط ثم نرسم مستطيلات في الاتجاه الموجب تبين الزيادة التي حدثت في الظواهر التي ازدادت قيمها ومستطيلات أخرى في الاتجاه السالب تبين النقص الذي حدث في الظواهر التي نقصت قيمها . ويجب أن نلاحظ أن أطوال المستطيلات لا تتناسب مع قيم هذه الظواهر وانما مع مقدار الزيادة أو النقص الذي حدث فيها ومن الواضح أنه يجب التفرقة بين المستطيلات التي في الاتجاه الموجب وتلك التي في الاتجاه السالب بالتلوين أو باستعمال التظليل المختلف .

ويمكن أيضاً استخدام شكل الأعمدة في توضيح التركيب الجنسي والعمري للسكان في احدى الدول حيث توضح فئات الاعمار بين عمودين في منتصف الرسم، ثم نرسم مستطيلات تبين عدد الذكور أمام كل فئة من الفئات في الجهة اليمنى من الرسم ومستطيلات تبين عدد الأناث أمام كل فئة من فئات الاعمار في الجهة اليسرى من الرسم . وفي هذه الحالة يكون المقياس المدرج على المحور الافقي وعلى أساسه تتناسب أطوال المستطيلات مع عدد الذكور وعدد الاناث في كل فئة من فئات الأعمار . ويسمى هذا الرسم بهرم التركيب العمري للسكان وله فوائد كثيرة حيث أنه يوضح النمط السكاني الذي يمر به المجتمع ، فاذا كان نمطاً بدائياً فان قاعدة الهرم تكون واسعة جداً بالنسبة للمدرجات الاخرى التي فوقها حيث تكون معدلات المواليد في هذه الدول مرتفعة جداً ، وكذلك تتقلص درجات الهرم تقلصاً سريعاً مما يدل على ارتفاع معدلات الوفاة . اما اذا كان نمطاً انتقالياً فان قاعدة الهرم تكون واسعة كذلك دلالة

على ارتفاع معدلات المواليد ، الا أن درجات الهرم لا تتقلص سريعاً دلالةعلى انخفاض معدلات الوفيات (السكان في الدول الشرقية) . أما اذا كان نمطه يدل على النضوج السكاني فان قاعدة الهرم تكون ضيقة نسبياً دلالة علىانخفاض معدلات المواليد كثيراً ، كما تكون درجات الهرم متقاربة جـداً دلالة على انخفاض معدلات انوفيات الى أدناهـــا وبذلك يرتفع الهرم كثيراً دلاله على ارتفاع العمر المتوقع للطفل المولود . كذلك يمكن أن يدلنا الهرم على ما اذا كان السكان يتأثرون في نموهم بالعوامل الطبيعية فقط (المواليد والوفيات)أو بعوامل شاذة كذلك (الهجرة والحروب .) حيث أن الدول التي يهاجر اليها السكان ينبعج الهرم الخاص بها في فئات العمر الخاصة بالشباب من ناحية الذكور والعكس اذا كان يهاجر منها السكان بتقلص الهرم تقلصاً شاذاً في هذه الفئات . كذلك يمكن أن يدلنا الهرم على أثر الحرب على النمو السكاني حيث نلاحظ ان مستطيلات الذكور تكون في الغالب أقصر من مستطيلات الاناث ، وكذلك نلاحظ تقلص المستطيلات الخاصــة بالمواليد بسبب تغيب الذكور أثناء الحرب .

ونلاحظ انه في كل هذه الأنواع من الرسومات التوضيحية يمكن أن نلجأ اليها سواء كانت الارقام التي لدينا ارقاماً مطلقة أو أرقاماً نسبية ، فبالنسبة لهرم التركيب العمري مثلاً يمكن أن تدل المستطيلات على عدد كل من الذكور والاناث في كل فئة عمر ويمكن أن تدل علىنسبة الذكور أو الاناث في كل فئة الى المجموع الكلي للسكان والفارق الوحيد ان المقياس المدرج في الحالة الاولى يظهر أرقاماً مطلقة أما في الحالة الثانية فيظهر أرقاماً نسبية .



NET NATIONAL PRODUCT AND PER CAPITA INCOME
(CURRENT FACTOR COST)

000 MILLION RS.

PER CAPITA INCOME

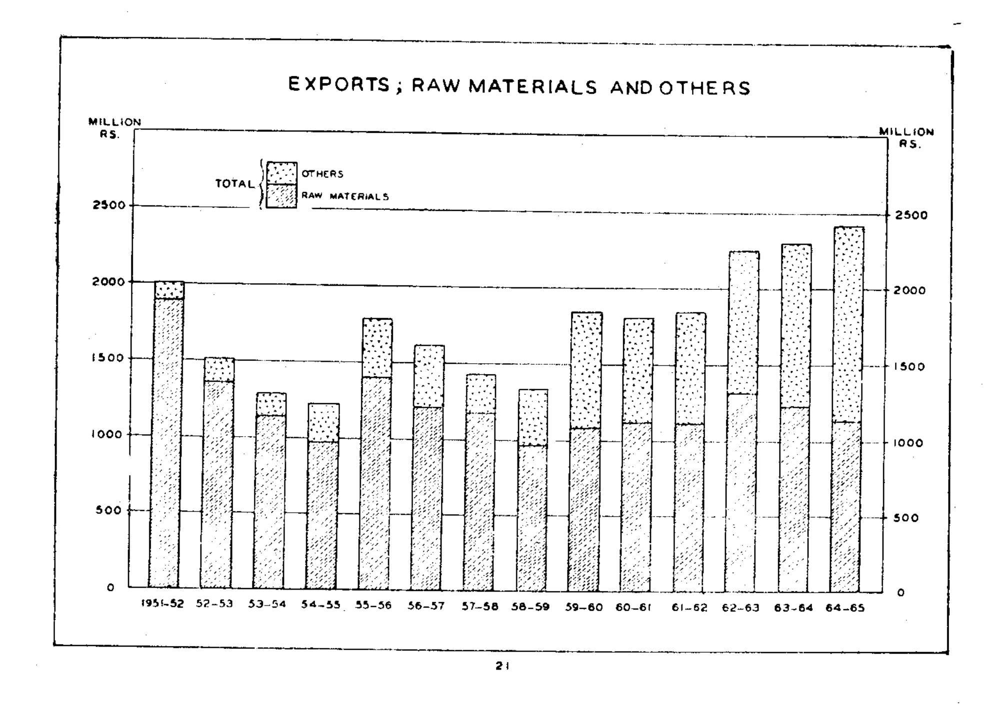
RS.

350 300 250 200

51-52 53-54 55-56 57-58 59-60 61-62 63-64

35 30 25 20 15 10

1951-52 52-53 53-54 54-55 55-56 56-57 57-58 58-59 59-60 60-61 61-62 62-63 63-64 64-65



شكل الدائرة :

يستعمل شكل الدائرة في عرض التوزيع النسبي للأوجه المختلفة لظاهرة معينة في فترة زمنية واحدة ، على أن تكرن هذه الأوجه مكونة للمجموع الكلي لهذه الظاهرة ، مثلا توزيع المساحة الكلية للاراضي الزراعية في الدولة تبعاً لأوجه استخدامها المختلفة ، والانواع المختلفة للقروض التي منحها أحد المصارف ... الخ . ولرسم هذا الشكل التوضيحي نرسم دائرة ذات قطر مناسب (ليس كبيراً جداً ولا صغيراً جداً) ونقسمها الى قطاعات تتناسب مساحتها مع نسبة كل جزء الى المجموع الكلي . ولاجراء ذلك نقسم الزاوية

المركزية والتي تساوي ٣٦٠° تبعاً لعدد الاجزاء التي تتكون منها الظاهرة وتبعاً لنسبة كل جزء من المجموع . فاذا فرضنا أن مجموع الاجزاء ٧٢٠ الف دونم وقيمة الجزء الاول ٣٤٦ الف دونم وقيمة الجزء الثاني ١٠٨ الف دونم وقيمة الجزء الثالث ٣٦٦ الف دونم يكون تقسيم الزوايا كالآتي :

$$٣٦٠ \times \frac{٣٤٦}{٧٢٠} = ٣٦٠ \times ٫٤٨ = ١٧٢٫٨^\circ$$

$$٣٦٠ \times \frac{١٠٨}{٧٢٠} = ٣٦٠ \times ٫١٥ = ٥٤٫٠^\circ$$

$$٣٦٠ \times \frac{٢٦٦}{٧٢٠} = ٣٦٠ \times ٫٢٧ = ١٢٣٫٢^\circ$$

وعلى أساس هذه الزوايا ١٧٢٫٨ ، ٥٤ ، ١٢٣٫٢ نقسم مساحة الدائرة الى القطاعات الثلاث ، ويمكن تظليل هذه القطاعات أو تلوينها للتمييز بينها ويكتب داخل كل قطاع اسم الجزء الذي يمثله . ويمكن أن يكتب على محيط الدائرة وخارج كل قطاع النسبة التي يمثلها من المجموع الكلي .

ومن الواضح انه لا يمكن استعمال شكل الدائرة في مقارنة التوزيع النسبي لظاهرة معينة في سنتين أو أكثر حيث لا يتضح من الرسم المقارنة بين قيم الاقسام المتناظرة خاصة إذا كانت قيمها متقاربة في السنتين .

السلسلة الزمنية :

سبق ان ذكرنا ان شكل الأعمدة لا يتفق مع البيانات الخاصة بعدد كبير من السنوات حيث ان الرسم يظهر مكتظاً بالأعمدة فيفقد بذلك رونقه ، لذلك يكون من الافضل في هذه الحالة رسم سلسلة زمنية خاصة اذا كانت البيانات تمثل متغيرات متصلة أي يكون تغيرها مع الزمن تغيراً متصلاً ، والسلسلة

الزمنية تساعد بذلك في إظهار الاتجاه العام للظاهرة موضوع البحث والتغير الذي يحدث في هذا الاتجاه العام من فترة زمنية الى أخرى .

وفي هذا النوع من الرسم تظهر قيمة الظاهرة في كل فترة بنقطة معينة على ارتفاع يمثل هذه القيمة ، ثم نصل هذه النقاط ببعضها حسب تسلسلها الزمني. وإذا كنا نريد أن نوضح الاتجاه العام للظاهرة نرسم خطاً أو منحنى يتوسط النقاط التي حددناها مقدماً خير توسط اي بحيث يكون مجموع انحرافات النقط عنه يساوي صفراً تقريباً ، إذ اننا في توضيح الاتجاه العـــام لا يهمنا إظهار التغير من عام إلى آخر وإنما الشكل العام لاتجاهها سواء كان مستقيما صاعداً أو هابطاً أو منحني محدباً إلى أعلى أو مقعّرا الى أسفل أو أي شكل آخر من الاشكال الرياضية التي تمثلها المعادلات المختلفة ، الدرجة الأولى أو الثانية أو الثالثة ... الخ .

ويمكن أن يحتوي الرسم على أكثر من سلسلة زمنية ،وفي هذه الحالة يجب توضيح كل منها بلون مختلف أو بأية طريقة مختلفة (خـــط متصل أو خط مقطع ... الخ) حتى نبرز اتجاه كل منها . ويحدث ذلك عندما نريد أن نظهر التطور الزمني في أكثر من ظاهرة واحدة ، على انه يحسن ان لا يظهر في الرسم أكثر من خمس خطوط بيانية ، إذ لو زاد العدد عن ذلك فإن تقاطعها أثناء اتجاهها يجعلها معقدة غير واضحة فتضيع الفائدة منها . وفي بعض الأحيان يكون الفراغ بين سلسلتين معبراً عن ظاهرة معينة اخرى نريد إبرازها في الرسم وبذلك نعمل على تظليلـــها بحيث لا يطمس التظليل معالم اتجاه السلسلتين اللتين تحدان الفراغ ، فعند رسم سلسلة زمنية للمواليد وأخرى للوفيات يكون الفراغ بينهما معبراً عن الزياة الطبيعية في السكان خلال الفترة موضوع البحث .

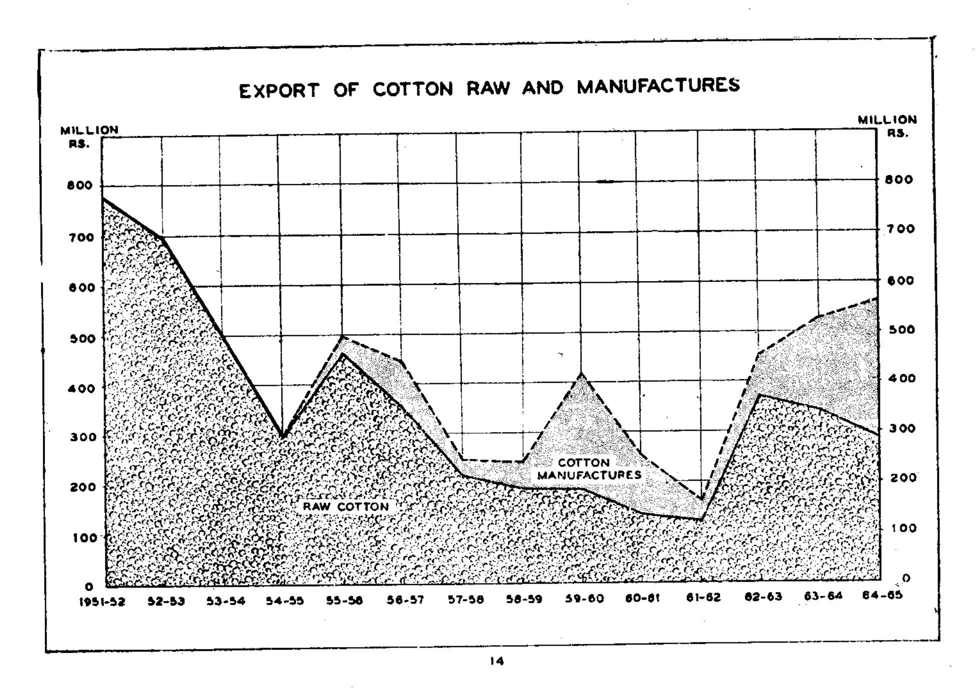
وفي بعض الاحيـــان يكون لدينا ظاهرة معينة مقسمة الى أجزاء في عدة سنوات ونريد ابراز التقسيم بجانب ابراز التغير الذي حدث فيه خلال

هذه المدة . لا نستطيع أن نوضح ذلك بشكل الدائرة حيث انه سبق أن ذكرنا ان التقسيم النسبي في عدة سنوات لا يمكن أن تظهره الدوائر الخاصة بالسنوات المختلفة . ومن الممكن توضيح هذه البيانات بأعمدة مجزأة للسنوات المختلفة التي يتضمنها البحث ، إلا ان كثرة الأعمدة كما ذكرنا يضيع على من يطلع على الرسم ملاحظة التغير الذي يصيب الأجزاء ، من عمود إلى آخر ، أي من فترة الى أخرى . لذلك يكون من الأفضل رسم ما نسميه بشكل القطاعات ، وهو في الواقع لا يختلف عن شكل السلسلة الزمنية إلا في حقيقة هامة وهي ان الخط البياني نفسه لا يمثل اتجاه الجزء الأول خلال سنوات البحث وإنما الذي يمثله هو الفراغ المحصور بين كل خط بياني وخط آخر ولذلك نسميه بشكل القطاعات . ولرسم هذا الشكل نرسم أولاً خطاً بيانياً يمثل اتجاه الجزء الأول من الظاهرة خلال سنوات البحث ثم خطاً بيانياً آخر يمثل اتجاه مجموع الجزئين الأول والثاني (وبذلك يكون الفراغ بينه وبين الخط الأول معبراً عن الجزء الثاني) ثم خطاً بيانياً ثالثاً يمثل اتجاه مجموع الثلاث أجزاء سوياً (وبذلك يكون الفراغ بينه وبين الخط الثاني معبراً عن اتجاه الجزء الثالث) . وهكذا حتى نصل الى آخر الخط الذي يمثل في الواقع اتجاه المجموع الكلي للظاهرة خلال سنوات البحث . بعد ذلك نظلل كل قطاع أو نلونه بلون مختلف ونكتب اسمه في داخله . والرسم بهذه الطريقة يوضح التركيب الجزئي للظاهرة في كل سنة وكذلك التغير في هذا التركيب من عام الى آخر .

كذلك يمكن ان يكون لدينا القيم الخاصة بظاهرة معينة في عدة فترات زمنية ونريد ان نوضح التغير الذي حدث في هذه القيم الذي يكون في بعض الاحيان موجباً وفي أحيان أخرى سالباً(مثلاً التغير في الأرقام القياسية حيث يمكن أن يزيد الرقم عن ١٠٠ وبذلك يكون هناك تغير موجب ويمكن أن تنقص القيمة عن ١٠٠ وبذلك يكون هناك تغير سالب). والرسم الذي يوضح هذه الحالة لا يختلف كثيراً عن الرسم العادي للسلسلة الزمنية والفرق الوحيد

بينهما هو في تقسيم الرسم الى جزئين يوضح الجزء الأعلى منه التغيرات الموجبة والجزء الأسفل يوضح التغيرات السالبة . ثم نحدد موضع النقط في الرسم حسب قيمة التغير وحسب إذا كان موجباً أو سالباً ثم نوصل النقط ببعضها حسب تسلسلها الزمني من فترة الى أخرى

ويمكن أن يكون لدينا قيم خاصة بعدة ظواهر في عدة فترات زمنية بحيث تكون وحدة القياس لهذه القيم مختلفة ، كما تختلف الأرقام عن بعضها اختلافاً كبيراً حيث يكون بعضها صغيراً جداً والبعض الآخر كبيراً جداً ، وبذلك نقع في مشكلة تحديد مقياس الرسم الذي يتفق مع هذا النوع من الأرقام . والمثل على ذلك بيانات عن عدد المؤسسات وعـــدد المشتغلين فيها ومتوسط الأجر للمشتغل الواحد في عدة سنوات ، ونريد أن نوضحها بحيث يمكن أن يظهر في الرسم مقارنة التغير الذي حدث فيها خلال المدة موضوع البحث ، والطريقة لمعالجة هذه الحالة هي في تحويل الأرقام المطلقة إلى أرقام نسبية على اساس الفترة الأولى وبذلك تصبح وحدة الأرقام متشابهة (٪) ، كما تصبح متقاربة من بعضها . ثم ننفذ الرسم على أساس هذه الأرقام النسبية وسوف نلاحظ من الرسم انها جميعاً تبدأ من نقطة واحدة (١٠٠) ثم تتخذ اتجاهاً يتفق مع التغير الذي حـــدث فيها وبذلك تظهر المقارنة فيما بينها ويجب ان نلاحظ كتابة عنوان الرسم كالآتي – التغير النسبي في عـــدد المؤسسات والمشتغلين فيها ومتوسط الأجر في المدة من عام – الى عام –.



خرائط الدبابيس :

تستخدم هذه الخرائط لتوضيح البيانات التي تطرأ عليها تغيرات من وقت إلى آخر ، مثل انتشار وباء معين خلال فترة زمنية معينة وانتشار الجراد وتحركات الجيوش . وفي هذا النوع من الرسم يمكن أن يعبر الدبوس الواحد عن وحدة وحدة أو عن عدد معين من الوحدات ، وتوضع هذه الدبابيس على الخرائط حسب المكان الذي ظهرت فيه هذه الحالة أو الحالات ، وهكذا كلما ظهرت حالة أو عدد معين من الحالات يوضع الدبوس على الخريطة في المكان المعين ، وبذلك يستطيع المعنيون بالامر تتبع ما يحدث في الظاهرة من وقت إلى آخر . ويمكن استعمال دبابيس بالوان مختلفة أو باشكال هندسية مختلفة لتوضيح ظواهر مختلفة على الخريطة . ومن الواضح أن مثل هذه الخرائط لا يقصد منها العرض في النشرات الاحصائية وانما العرض المكتبي لاطلاع المسئولين .

كذلك يمكن استخدام الخرائط المنقطة لتوضيح البيانات الخاصة بظاهرة معينة في أجزاء الدولة المختلفة ، حيث توضع نقط مستديرة تمثل كل منها عدداً معيناً من البيان الخاص بالظاهرة موضوع العرض ، فلو رسمنا خريطة للبنان وقسمناها إلى محافظات فانه يمكن التعبير مثلاً عن انتاج للفواكه في كل محافظة بمثل هذا النوع من الخرائط . هذه الخرائط صالحة للطبع في نشرات حيث انها ليست مثل خرائط الدبابيس إذ المقصود منها اظهار توزيع الظاهرة على أجزاء الدولة في وقت معين وليس إظهار التغير الذي يحدث في انتشار الظاهرة من وقت الى آخر .

الرسم البياني اللوغاريتمي :

يحتاج الاحصائي أحياناً الى توضيح معدل التغير في الظواهر وليس التغير المطلق في قيمها . ولتفسير ذلك نفترض ان لدينا عدد السكان في مدينتين في تعدادين ونريد أن نقارن التغير في هذه الأعداد لنعرف أي المدينتين تنمو بسرعة أكبر . نستطيع أن نحدد مقدار التغير في عدد سكان كل مدينة بطرح عدد سكانها في التعدادين ثم نقارن ناتج الطرح ، غير انه لكي تكون المقارنة سليمة لا بد أن ننسب عدد سكان كل مدينة الى عددها في التعداد الأول حيث ان ذلك هو الذي يوصلنا فعلاً الى المقارنة السليمة بينهما حيث أن هدفنا هو في الواقع معرفة أي المدينتين تنمو بمعدل أكبر ؟ فاذا فرضنا ان عدد سكان المدينة الأولى هو ١٠ ملايين نسمة في تعداد ١٩٥٠ وأصبح ١٢ مليون نسمة في تعداد ١٩٦٠ ، وان عدد السكان المدينة (ب) هو ٥ مليون نسمة في تعداد ١٩٥٠ وأصبح ٧ مليون نسمة في تعداد ١٩٦٠ ، هل نستطيع أن نستنتج من ذلك ان معدل النمو في سكان المدينتين متساو ؟ ان هذه النتيجة تكون في الواقع نتيجة مضللة حيث ان سكان المدينة (أ) قد زادوا بنسبة ٢٠ ٪ بينما زاد سكان المدينة (ب) بنسبة ٤٠ ٪ ، لذلك لا نستخدم الرسم البياني العادي في توضيح معدلات التغير حيث أن

هـــذا الرسم لا يستطيع أن يظهر الا التغيرات المطلقة والرسم البياني اللوغاريتمي هو الوسيلة لتوضيح معدلات التغير حيث أن الفكرة الأساسية لهذا الرسم هي أنه بينما تكون المسافات المتساوية في الرسم البياني العادي ممثلة لقيم مطلقة متساوية تكون في الرسم البياني اللوغاريتمي ممثلة لمعدلات متساوية ولذلك نجد ان مقياس التدرج في الرسم البياني العادي يظهر الأرقام على أساس أنها تكون متوالية عديدة مثلا صفر ، ٥٠، ١٠٠ ، ١٥٠ ، ٢٠٠ ، وهكذا، أما في الرسم البياني اللوغاريتمي فتظهر الارقام على أساس أنها تكون متوالية هندسية ، ١٠٠ ، ١٠٠٠ ، ١٠٠٠٠ ، وهكذا وبذلك بينما يكون كل ١سم في الرسم البياني العادي ممثلا لخمسين وحدة مثلا ، يكون كل ١ سم في الرسم البياني اللوغاريتمي ممثلا لعشرة أضعاف .

تقوم فكرة الرسم البياني اللوغاريتمي على أساس الحقيقة الهامة وهي انه اذا كان لدينا ارقام تكون متوالية هندسية فان لوغاريتمات هـــذه الأرقام تكون متوالية عددية أى يكون الفرق المطلق بين أعدادها متساوياً مثلا :

الاعداد	١٠٠	١٥٠	٢٢٥	٣٣٧,٥	٥٠٦,٢٥
لوغاريتماتها	٢	٢,١٧٦	٢,٣٥٢	٢,٥١٨	٢,٧٠٤

ولاثبات ذلك نفترض ان لدينا القيم $س_1$ ، $س_2$ ، $س_3$ ، $س_4$ ، تكون متوالية هندسية :

$$\frac{س_2}{س_1} = \frac{س_3}{س_2} = \frac{س_4}{س_3}$$

$$لو \frac{س_2}{س_1} = لو \frac{س_3}{س_2} = لو \frac{س_4}{س_3}$$

$$لو\ س_2 - لو\ س_1 = لو\ س_3 - لو\ س_2 = لو\ س_4 - لو\ س_3$$

أي أن لوغارتيمات $س_1$ ، $س_2$ ، $س_3$ ، $س_4$ (وهي تكون هندسية كما قدمنا) تكون متوالية عددية .

وبذلك يمكننا ان نستفيد من هذه القاعدة في تصميم رسم بياني يوضح معدلات التغير في البيانات التي لدينا ، وذلك بدلا من ان نظهر القيم الفعلية لهذه البيانات (وهو يظهر تغيرها المطلق) نظهر لوغاريتمات هـذه الاعداد وبذلك يتضح تغيرها النسبي، ومن الواضح ان الظاهرة اذا كانت تتغير بمعدل ثابت فان لوغارتيمات القيم الخاصة بها سوف تتغير بمقدار ثابت، واذا كانت قيم الظاهرة تتغير بمعدلات متزايدة كان لوغاريتماتها تتغير بمقادير متزايدة، والعكس اذا كانت قيم الظاهرة تتغير بمعدلات متناقصة فان لوغاريتماتها تتغير بمقادير متناقصة.

وبذلك يتضح لنا انه يمكن ان نوضح معدل التغير في ظاهرة مـــا على القياس العادي بشرط الا نرصد قيم الظاهرة نفسها بل نرصد لوغاريتمات هذه القيم . ولتنفيذ الرسم بهذه الطريقة نضع القيم على المحور الرأسي على ابعاد تتناسب مع الفرق بين لوغاريتماتها . وعنـــد تقسيم المحور الرأسي تقسيما لوغاريتميا نضع على نقطة الأصل أي قيمة موجبة بخلاف الصفر (إذا وضعنا الرقم ١ يكون لوغاريتمه = صفر) ، ثم نأخذ على المحور بعداً يساوي لوغاريتم ٢ ونكتب عليه ٢ . ،حيث أن لوغاريتم ٢ = ٣٠١ ,٠ فانـنـــا نكتب القيمة ٢ على بعد ٣٠١و٠ سم من نقطة الأصل، إلا أن ذلك يجعل الرسم من الصفر بحيث يتعذر تحديد النقط عليه فيما بعد . لذلك يكون من الأفضل أن نأخذ مسافة ٣٫٠١ سنتمتر مثلاً لتعبر عن القيمة ٢ ثم نأخذ من نقطة الأصل مسافة تساوي لوغاريتم ٣ بنفس الوحدات التي استعملناها للتعبير عن لوغاريتم ٢ ، ونكتب على هذه المسافة الثانية الرقم ٣ وهكذا بالنسبة لباقي القيم .

بعد ذالك نستطيع أن نضع أمام التقسيات التي قسمناها أي سسلة من الأرقام تكون متناسبة مع الأعداد الطبيعيه من ١ إلى ١٠ فنستطيع مثلاً أن نبدأ بالعدد ١٠٠ بدلاً من ١ بحيث نضع ٢٠٠ في مكان ٢ و ٣٠٠ في مكان ٣ وهكذا إلى أن نصل إلى الف ، كذلك يمكن أن نبدأ بالقيمة ١٥ فنكون باقي القيم ٣٠ ، ٤٥ ، ٦٠ وهكذا حتى تصل إلى ١٥٠ ، كذلك يمكن أن

نبدأ بالقيمة ٥٠ فيكون باقي القيم ١٠٠ ، ١٥٠ ، ٢٠٠ ، ٢٥٠ ، ٣٠٠ . حتى نصل إلى ٥٠٠ . ومن الواضح اننا نختار القيم التي تتفق مع القيم التي نريد توضيح معدل تغيرها بالرسم البياني . والمهم أن نلاحظ انه مهما غيرنا من القيم فلا بد أن تتناسب لوغاريتماتها مع لوغاريتمات الاعداد الطبيعة من ١ الى ١٠ . ويتضح السبب في ذلك مما يأتي : —

$$\text{لو } ١٠٠ - \text{لو } ٥٠ = \text{لو } \frac{١٠٠}{٥٠} = \text{لو } ٢ - \text{لو } ١$$

$$\text{لو } ٣٦ - \text{لو } ١٨ = \text{لو } \frac{٣٦}{١٨} = \text{لو } ٢ - \text{لو } ١$$

$$\text{لو } ٩٠ - \text{لو } ٤٥ = \text{لو } \frac{٩٠}{٤٥} = \text{لو } ٢ - \text{لو } ١$$

وبذلك تكون المسافة بين لو ١٥٠٠ولو ٧٥٠ هي نفس المسافة بين لو ٤٨ ولو ٢٤ هي نفس بين لو ٢ ولو ١ .

بعد تحديد المقياس اللوغاريتمي على المحور الرأسي نستخرج لوغاريتمات القيم التي نريد توضيحها ونرصدها في الرسم تبعاً للمقياس الذي حددناه . وحتى تتفق اللوغاريتمات مع لوغاريتمات الأعداد الطبيعة نقسم القيمة التي نريد رصدها في الرسم على أول قيمة بدأنا بها على المحور الرأسي ثم نستخرج لوغاريتم ناتج القسمة ، فادا كنا قد بدأنا المحور بالقيمة ٥٠ ونريد رصيد القيمة ٥٥ على أساس المقياس اللوغاريتمي ، نقسم أولاً ٥٥ على ٥٠ = ١٫١ ثم نستخرج لوغاريتم هذه القيمة ونرصده تبعاً للمقياس الذي اتخذناه على المحور الرأسي معبراً عن اللوغاريتمات .

لاحظنا أن القيم التي يمكن رصدها تبعاً للمقياس اللوغاريتمي الذي شرحناه تتغير بنسبة تتراوح من ١ إلى ١٠ ، مثلاً ٥٠ الى ٥٠٠ ، أو ٧٠ إلى ٧٠٠ . هذا المدى يسمى دورة لوغاريمية . لكن ما العمل إذا كانت القيم تتغير في

مدى أوسع من ذلك مثلا من ٥٠ إلى ٢٠٠٠ مثلا ؟ نستطيع أن نكمل التقسيم بدوره ثانية ودورة ثالثة وهكذا – تكون الدورة الأولى من ٥٠ إلى ٥٠٠ والدورة الثانية من القيمة ٥٠٠ وتنتهي بالقيمة ٥٠٠٠ وتبدأ الثالثة بالقيمة ٥٠٠٠ وتنتهي بالقيمة ٥٠٠٠٠ وهكذا . ومن الواضح أن كل دورة يتفق تقسيمها مع تقسيم لوغاريتمات الأعداد الطبيعة من ١ إلى ١٠ . كذلك يتضح لنا أن الرسم اللوغاريتمي يمكن أن يتسع لقيم تبدأ صغيرة جداً وتنتهي كبيرة جداً .

عندما يكون التقسيم اللوغاريتمى على المحور الرأسى فقط أي أن المحور الأفقي يكون تقسيمه عادياً يسمى الرسم نصف لوغاريتمي ، أما إذا كان التقسيم اللوغاريتمي على كلا المحورين يكون الرسم لوغاريتمي كامل .

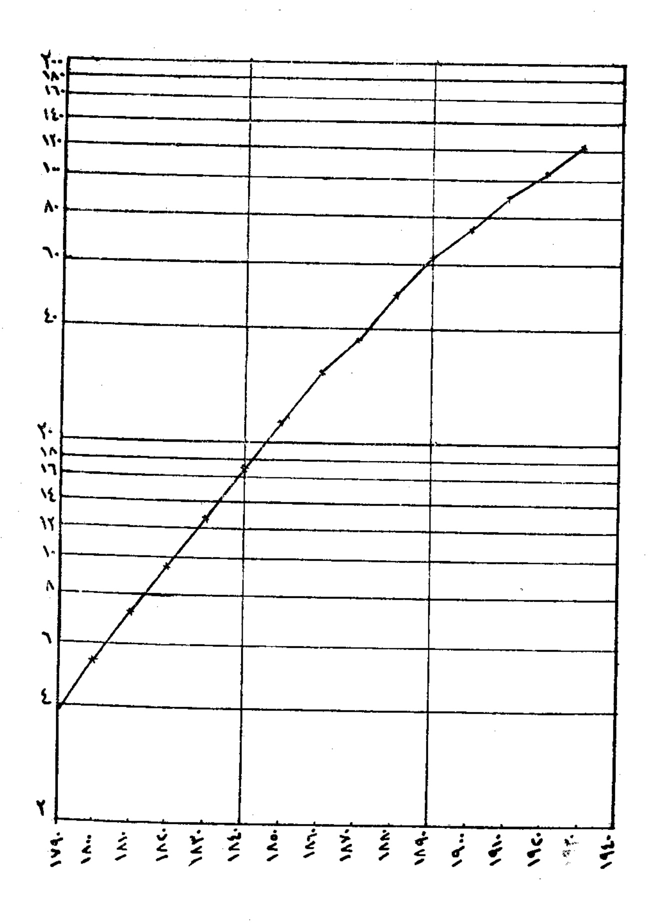
بعد رصد النقط على الرسم البياني النصف اللوغاريتمي وتوصيلها سويا بخط بياني نستطيع أن نتعرف على نوع التغير في الظاهرة من ميل هذا الخط، فإذا كانت الظاهرة تتغر بمعدل ثابت يظهر ذلك في الرسم من وقوع جميع النقط على الخط المستقيم الذي يصل بينها كما ان درجة ميل هذا الخط توضح لنا ما إذا كان التغير بمعدل متزايد أو معدل متناقص . كذلك إدا رسمنا خطين بيانيين لظاهرتين مختلفتين على رسم نصف لوغاريتمي نستطيع أن نقارن معدل التغير فيهما تبعاً لميل الخطين، فاذا توازى الخطان يدل ذلك على تساوي معدل التغير في الظاهرتين ، أما إذا اختلف ميلها يدل ذلك على اختلاف معدل التغير فيهما .

وهناك تطبيقات عملية كثيرة للرسم البياني النصف اللوغاريتمي حيث يمكن أن نلجأ اليه عندما تكون الظاهرتان مقاستان بوحدات مختلفة أو عندما تكون قيمها متباعدة كثيراً حيث تكون قيم احداها صغيرة جداً بينما تكون قيم الأخرى كبيرة جداً . كذلك يلجأ الى الرسم اللوغاريتمي عندما تتغير قيم الظاهرة في مدىواسع جداً حيث تبدأ صغيرة وتنتهي بقيم كبيرة جداً .

وفي المجال الاقتصادي يستخدم الرسم اللوغاريتمي لحساب مرونة الطلب حيث نرسم منحنى الطلب على رسم لوغاريتمي كامل ثم نرسم مماساً لمنحنى

الطلب عند النقطة المقابلة للثمن الذي نريد تقدير المرونة عنده ثم نحسب ميل هذا المماس على محور الثمن حيث أن هذا الميل سوف يوضح لنا ناتج قسمة التغير النسبي في الطلب على التغير النسبي في الثمن . كذلك نستخدم الرسم اللوغاريتمي في توضيح قانون باريتو الذي ينص على انه كلما زادت قيمة الدخل بنسبة معينة كلما نقص عدد الأشخاص الذين يحصلون على هذا الدخل أو أكثر بنفس النسبة، وقد أثبتت الاحصاءات أن توزيع الدخل في بعض الدول يتمشى مع هذا القانون . فاذا أخذ مقياس لوغاريتمي للدخول على المحور الافقي ومقياس لوغاريتمي آخر على المحور الرأسي لتوضيح عدد الأشخاص الذين يحصلون على هذه الدخول ورصدنا النقط التي تمثل قيم الدخول والأشخاص الذين يحصلون عليه نجد انها تقع تقريباً على خط مستقيم، الأمر الذي يدل على ان نسبة التغير في الدخل تساوي نسبة التغير في عدد الأشخاص .

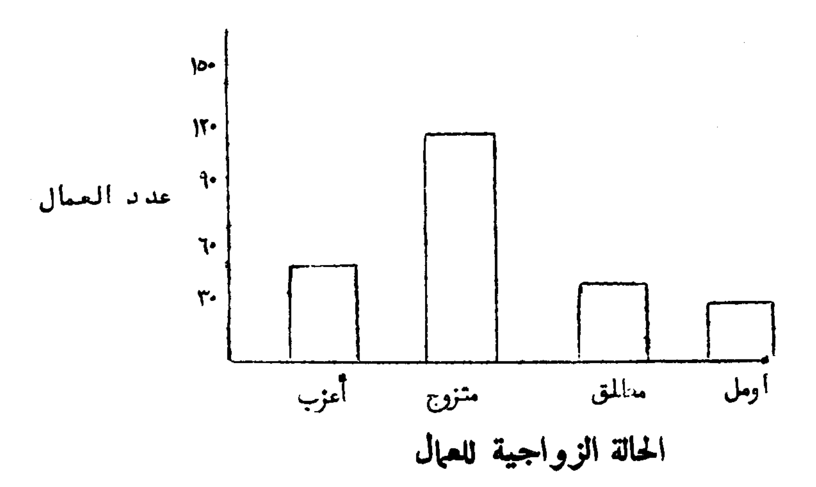
كذلك نستخدم الرسم النصف اللوغاريتمي في توضيح معدل التغير في الظواهر مع الزمن مثل نمو السكان حيث أن ميل الخط الذي يمثل هذا النمو يدل على ما اذا كان السكان يزيدون بمعدل ثابت أو بمعدل متزايد عندما يكون الخط في شكل منحنى يصعد بسرعة من اليسار الى اليمين أو بمعدل متناقص عندما يكون الخط في شكل منحنى يصعد ببطء .



التمثيل البياني للتوزيعات التكرارية :

هناك أربعة وسائل أساسية لتمثيل أي توزيع تكراري في شكل بياني :

(١) طريقة الأشرطة البيانية (الأعمدة) ، وتتميز هذه الطريقة بصلاحيتها بنوع خاص لتمثيل التوزيعات التكرارية غير الرقمية وكذلك التوزيعات ذات القيم الوثابة (الغير متصلة) ، وتقوم هذه الطريقة على أساس تمثيل المتغير على المحور الافقي والتكرارات على المحور الرأسي فيكون الشكل البياني عبارة عن مجموعة من الأشرطة يمثل قاعدة كل منها المتغير وارتفاعه أو طوله تكرار هذه المتغير. خذ التوزيع التكراري غير الرقمي المشار اليه سابقاً والخاص بتوزيع العمال حسب حالتهم الزواجية فانة من الممكن تخصيص مسافات متساوية في العرض تبعد كل منها عن الأخرى بابعاد متساوية ثم اقامة أعمدة أو مستطيلات على كل من هذه المسافات تتناسب أطوالها وتكرار كل مجموعة كما هو مبين في الشكل التالي:

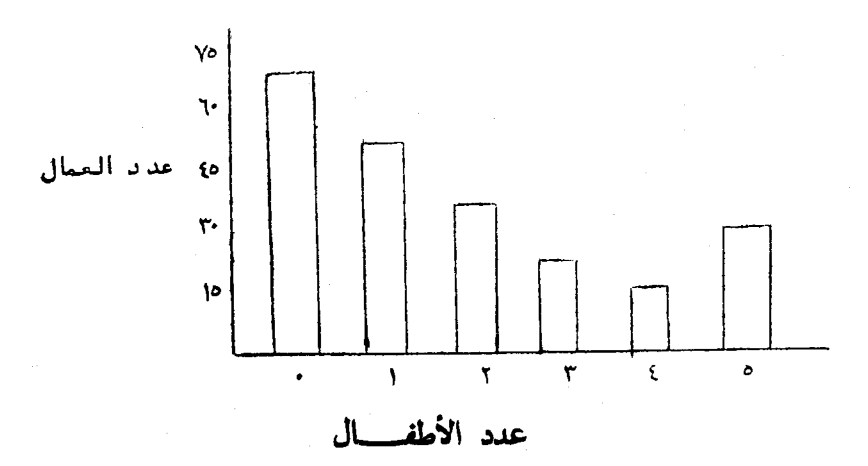


الحالة الزواجية للعمال

ولتوضيح تطبيق تلك الطريقة على التوزيعات التكرارية ذات المتغيرات الوثابة نأخذ تقسيم هؤلاء العمال حسب عدد أولاد كل منهم فنحصل على الجدول الآتي :

عدد الأطفال	٠	١	٢	٣	٤	٥ وأكثر	المجموع
عدد العمال	٧٧	٥٨	٤٠	٢٧	١٥	٣٣	٢٥٠

ويكون الشكل البياني كالآتي :



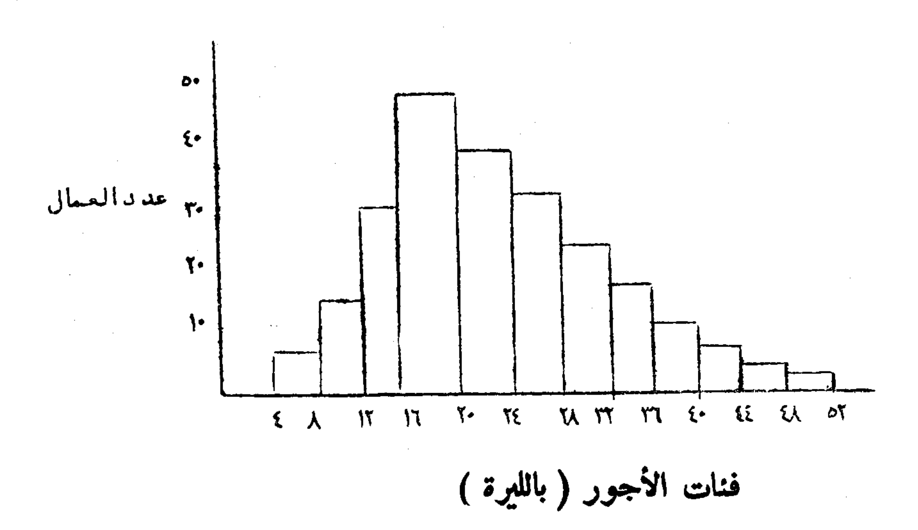
عدد الأطفال

٢- والطريقة الثانية وهي الأكثر شيوعاً – هي طريقة المدرج التكراري أو الهستوجرام (Histogram) وتنحصر تلك الطريقة في أخذ مسافات على المحور الأفقي تمثل أطوال فئات التوزيع واقامة مستطيلات ملتصقة تتناسب أطوال كل منها وتكرار كل فئة .

فلو أردنا تمثيل التوزيع التكراري للعمال السابقين حسب أجورهم فما علينا إلا ان نأخذ مسافات متساوية تمثل كل منها أربعة ليرات على المحور الافقي واقامة مستطيلات على تلك المسافات تمثل مساحة كل منها عدد العمال ذوي الأجر المحـدد في الفئة . ولمـــا كانت قواعد المستطيلات كلها متساوية (الفئات المتساوية)، لهذا نكتفي برسم مستطيلات يمثل طول كل منها التكرار المقابل للفئة ، فمثلاً إذا كان التوزيع كالآتي :

فئات الأجر	٤-	٨-	١٢-	١٦-	٢٠-	٢٤-	٢٨-	٣٣-	٣٦-	٤٠-	٤٤-	٤٨- ٥٢
عدد العمال	٦	١٤	٣١	٥٤	٤٣	٣٦	٢٧	١٩	١٠	٦	٣	١

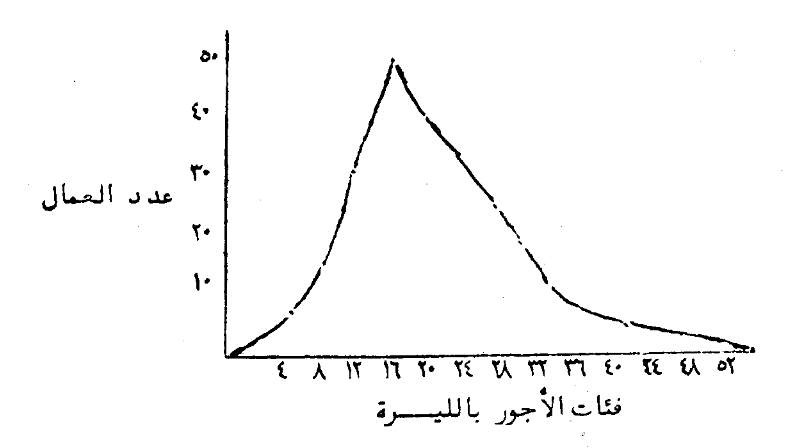
يكون المدرج التكراري كالآتي :



فئات الأجور (بالليرة)

ونلاحظ ان المساحة المحدودة بالمدرج التكراري هي مجموع التكرارات وذلك لأن مساحة كل مستطيل صغير في المدرج تساوي تكرار الفئة فاذا جمعنا هذه المستطيلات الصغيرة حصلنا على مجموع يمثل مجموع التكرارات هذا على أساس أن قواعد المستطيلات متساوية ولذلك يكون التعبير عنها بالرقم ١ .

٣) والطريقة الثالثة تسمى طريقة المضلع التكراري أو البوليجون (Polygon) وتنحصر تلك الطريقة في تحديد مسافات على المحور الأفقي يمثل كل منها فئات التوزيع المطلوب رسمه ثم وضع نقطة بأعلى مركز كل فئة بحيث يتناسب بعد النقطة عن مركز الفئة مع تكرارها ثم توصيل تلك النقط بخطوط متكسرة . والبوليجون بعبارة اخرى ما هو إلا مجموعة الخطوط المتكسرة التي توصل منتصف قمم المدرج التكراري ، وتطبيقاً لتلك الطريقة يمكن تمثيل التوزيع السابق كالآتي :



ولكي نقفل المدرج نفترض وجود فئة سابقة للتوزيع وتكرارها يساوي صفر وكذلك فئة لاحقة وتكرارها أيضاً يساوي صفر ثم نصل نهايتي المضلع بمركزي الفئتين . ونلاحظ ان المساحة المحدودة بالمضلع التكراري

تساوي المساحة المحدودة بالمدرج أي تساوي مجموع التكرارات وذلك لأن المضلع يضيف أجزاء الى مساحة المدرج ويستبعد أجزاء منها ، ومجموع الأجزاء المضافة يساوي الأجزاء المستبعدة وبذلك تبقى مساحة المدرج دون تغيير .

٤) والطريقة الرابعة هي طريقة المنحنى التكراري Frequency Curve وهي من أهم الطرق من الوجهة النظرية .

وتتلخص هذه الطريقة في تحديد أطوال الفئات المختلفة على المحور الأفقي ثم وضع نقط بأعلى مركز كل فئة بحيث يتناسب البعد بينها وبين المركز مع تكرار الفئة ثم توصل النقط بخط ممهد . وما المنحنى بعبارة أخرى الا المضلع التكراري نفسه ممهداً بحيث يتخذ شكل منحني .

والمساحة التي يحدها المنحنى لا تساوي مساحة المدرج لأن المنحنى يضيف أجزاء ويستبعد أجزاء أخرى لا تساويها .

ويجدر بنا الاشارة الى العلاقة بين المنحنيات والمدرجات التكرارية إذ ان المنحنى التكراري يعطي في الواقع الصورة العامة للعلاقة بين المتغير وتكراراته لا للتوزيع التكراري موضوع الدراسة فحسب بل للتوزيع التكراري العام الذي اشتق منه هذا التوزيع .

خذا المثل الآتي لتوضيح تلك الفكرة :

لو أخذنا مجموعة صغيرة من طلبة الجامعة العربية ببيروت وليكن عددهم ٢٠٠ طالباً كعينة ممثلة لأفراد المجتمع الأصلي وهو طلبة الجامعة بأجمعها : ثم سجلنا أعمار أفراد تلك العينة وعملنا توزيعاً تكرارياً لهذه الأعمار فان المدرج التكراري لهذا التوزيع يعطينا العلاقة بين السن والعدد أو بعبارة أخرى يعطينا قانون تغيّر السن للمجموعة . تصور اننا كبرنا عدد أفراد تلك العينة وجعلناه ١٠٠٠ طالب لتمثيل المجتمع الأصلي نفسه ثم عملنا توزيعاً تكرارياً لأعمار الألف طالب ورسمنا المدرج التكراري لهذا التوزيع .

واضح أن هذا المدرج التكراري سيعطينا نفس قانون التغير السابق على فرض ان كل من المجموعتين الصغيرة والكبيرة تمثل المجتمع الأصلي تمثيلاً صادقاً.

تصور خطوة أخرى اننا أخذنا جميع طلبة الجامعة بلا استثناء ورسمنا مدرجاً تكرارياً لتوزيع أعمارهم فان هـذا المدرج سيعطينا نفس قانون التغير السابق .

ما الفرق إذن بين المدرجات التكرارية الثلاث التي رسمناها ؟ لا فرق بينها إلا في كثرة تكرارات التوزيع الأخير بدرجة تسمح بطبيعة الحال من تقصير مدى الفئات حتى يمكن تصور ان النقط الممثلة لقمم المستطيلات المختلفـة تتقارب من بعضها البعض حتى تكون خطاً ممهداً هو المنحنى التكراري الذي كان يمكن الحصول على شكله العام بالتقريب من المدرج التكراري الأول .

نخرج من ذلك بحقيقة هامة ان المدرج التكراري يظهر قانون التغير بين أفراد عينة من مجتمع أصل بينما يظهر المنحنى التكراري قانون التغير لهـذا المجتمع نفسه .

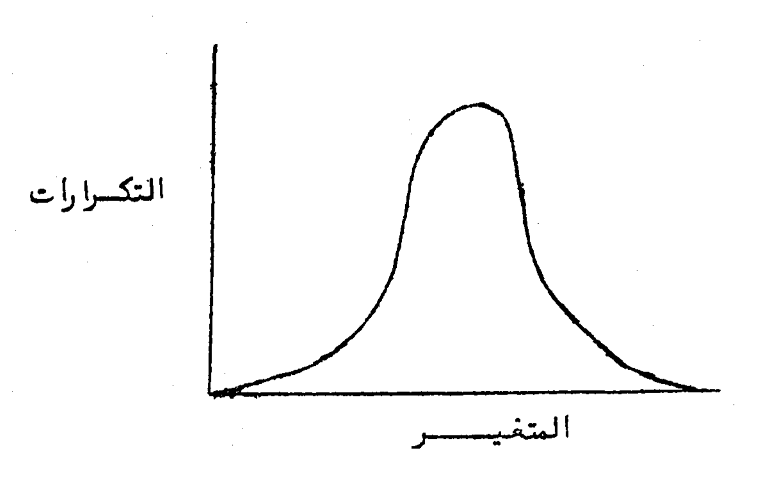
أشكال المنحنيات التكرارية :

خلصنا من دراستنا السابقـة بأن شكل المنحنى التكراري لأي توزيع يعطي فكرة صادقة عن قانون التغير فيه .

ومن المعروف ان التوزيعات التكرارية مهما اختلفت في طبيعتها فانهـا تنقسم إلى أنواع محدودة العدد واضحة المعالم ، إذ يمكن تقسيم جميع التوزيعات التكرارية وبالتالي منحنياتها الى قسمين رئيسيين الأول – توزيعات متماثـلة (symmetric) والثاني توزيعات غير متماثلة أو ملتوية (Asymmetric skewed)

والتوزيعات أو المنحنيات المتماثلة هي التي تكون العلاقة بين المتغير والتكرارات بحيث تتزايد التكرارات تدريجياً كلما زاد المتغير حتى تصل التكرارات الى قمة معينة تبدأ بعدها في النزول كلما زادت القيم أكثر من ذلك ، وبحيث يكون

الصعود إلى القمة بنفس السرعة التي تهبط بها التكرارات بعـــد القمة كما هو واضح في الشكل التالي. وواضح ان المنحني متماثل في صعوده وهبوطه حول قيمة معينة للمتغير ومن هنا جاءت التسمية (التماثل) .



وهذا المنحنى في شكله العـــام يشبه الناقوس ولذا يسميه البعض منحنى ناقوسي (Bell shaped Curve).وهذا النوع من التوزيعات من أهم الأنواع وخصوصاً في الدراسات الاحصائية النظرية . ويغلب حدوثه في التوزيعات التكرارية « الطبيعية » أي التي لا يؤثر العامل البشري على قانون التغير فيها ولذا يسميه البعض الآخر منحنى طبيعي (Normal) مثل توزيعات عـــلم الاحياء وكذلك في علم الأجناس وفي المسائل التي يؤثر عامل الحظ (الصدفة) وحده في حدوثها مثل التوزيعات التكرارية لعدد مرات ظهور وجه معـين للعملة عند رمي قطعة منها عدة مرات .

ولهـــذا المنحنى خواص احصائية ورياضية معينة سنحاول دراستها في الفصول القادمة .

وهنـــاك نوع آخر من المنحنيات المتماثـــلة وهو المنحنى النوني (U. shrped Curvs) وهذا النوع على عكس السابق في أن أكبر التكرارات

تقابل أصغر القيم وأكبرها ، وأصغر التكرارات تقابل تقاطع خط التماثل مع محور المتغير .

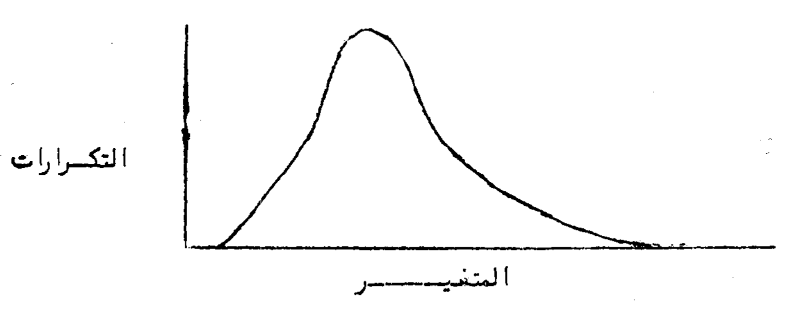
وأمثال هذه التكرارات قليلة الحدوث جداً والمثل الذي كثر ترديده بين الكتاب لهذا النوع من المنحنيات هو توزيع درجة تكاثف السحب في سماء جرينتش في سنة من السنين . ومن أمثلة هذا النوع من التوزيع التكراري نسبة البطالة بين عدد كبير من العمال بحسب أعمارهم اذ أن نسبة البطالة بين صغيري السن كبيرة ثم تتضاءل كلما كبر السن الى حد معيّن تبلغ عنده نهايتها الصغرى ثم ترتفع تدريجياً بعد ذلك كلما ارتفع سن العامل

ولا بد لنا أن نذكر بصراحة انه قل أن يتوفر في التوزيعات التكرارية العملية توزيعاً من هذا النوع يكون تاماً في تماثله ولكنه على أي حال يمكن اعتبار الكثير من تلك التوزيعات قريبة من التماثل بشيء من التجاوز .

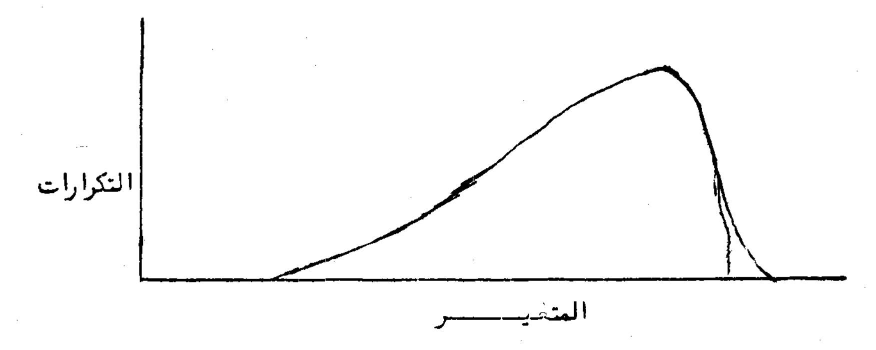
التوزيعات الملتوية :

أما المنحنيات الملتوية فهي اما معتدلة الالتواء أو حادة الالتواء . وهناك طائفتان من النوع الأول ، الأولى هي المنحنيات الموجبة الالتواء وهي ما تزايد التكرار فيها بسرعة كلما تزايدت قيمة المتغير حتى تصل الى قمتها ثم تنخفض التكرارات ببطء كلما تزايدت القيم بعد ذلك . ويستطيع القارىء أن يرى مثلاً من أمثال تلك التوزيعات في الشكل التالي وتسهيلا له الالتواء تحديد طبيعة هذا النوع من المنحنيات يجب أن يذكر أن المنحنى الموجب الالتواء هو ما كان « ذيله » الى اليمين . ويكثر حدوث مثل هذا التوزيع في المسائل الاجتماعية والاقتصادية . حيث تميل أغلب التكرارات الكبيرة الى ناحية القيم الصغيرة للمتغير . فمثلا لو أخذنا التوزيع التكراري للمتزوجين في بلد معين بحسب السن فان عددهم عند العمر ١٨ يكون قليلاً ثم يكبر عددهم كلما زاد العمر حتى يصل العدد أقصاه عند العمر ٢٢ ، ٢٣ سنة مثلاً ثم يبدأ

عددهم يتناقص لكل فئة من العمر ببطء حتى يصل العدد الى أقله عند العمر ٦٠ ، ٧٠ سنة .

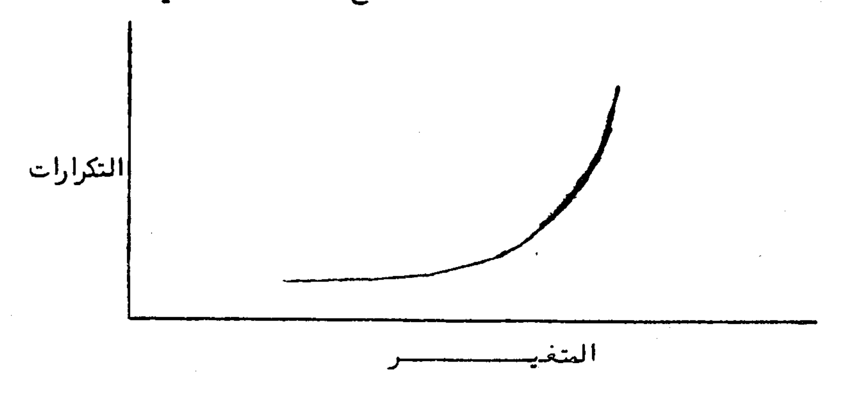


والطائفة الثانية من المنحنيات المعتدلة الالتواء هي التي تمتاز ببطء صعود تكراراتها كلما زاد المتغير حتى تصل التكرارات الى قمتها عند نقط معينة ثم تبدأ في النزول بسرعة كلما زاد المتغير بعد ذلك ويسمى هذا النوع منحنى ملتو التواء سالباً حيث يكون « ذيله » الى اليسار كما هو واضح في الشكل التالي . وهذا النوع من المنحنيات قل أن يصادفه الباحث وخصوصاً بيـــن الظواهر الاجتماعية والاقتصادية ولذلك فامثلتة في هذا الميدان قليلة ، وعلى الرغم من ذلك اذا عملنا توزيعا تكرارياً لأعمار المواليد بحسب أعمار الأمهات فان التوزيع التكراري الناتج يكون من هذا النوع . ويكـثر على العكس من ذلك أمثال هذا التوزيع في المسائل الخـــاصة ببعض المقاييس الجوية . فلو قمنا بعمل توزيع تكراري لبضع مئات من قراءات ضغط الهـــواء في مدينة معينة فان التوزيع الناتج يكون ملتوياً التواء سالباً .



المنحنيات حادة الالتواء :

هناك نوعان من هذه المنحنيات ، منحنيات تتزايد التكرارات فيها كلما كبرت قيمة المتغير حتى تصل التكرارات اقصاها عند نقطة معينة من المتغير ولا تعود الى الانخفاض بعد ذلك كما هو واضح من الشكل الآتي :



ويسمى هذا النوع من المنحنيات منحنى رائى للشبه بينه وبين حرف الراء العربية ، وكذلك اطلق عليه في الانجليزية (J. Shaped Curve) للشبه بينه وبين حرف J .

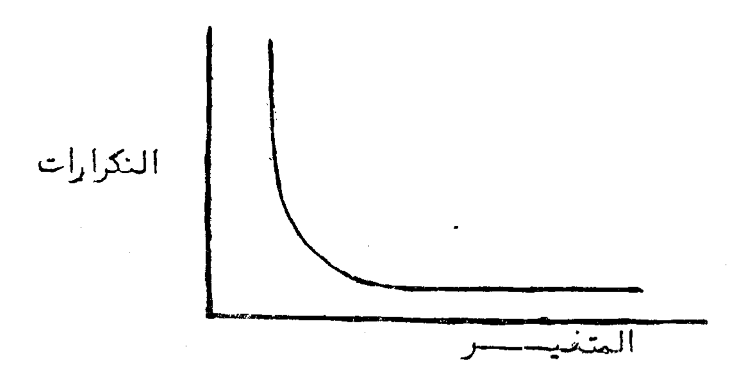
أما إذا كانت التكرارات كبيرة نوعا مـــا عند مبدأ هذا المنحنى ثم تناقصت بعد ذلك فنطلق على هذا النوع من المنحنيات اسم منحنى لامي تشبيهاً له بحرف اللام العربية ويلاحظ ان الكتاب الانجليز لا يفرقون بين النوعين اللامي والرائي فيستخدمون اسما واحداً لهما .

وهذا النوع من التوزيع التكراري نادراً الحدوث في المسائل الاجتماعية . ولقد وجد أن التوزيع التكراري للمصابين في حوادث الطرق بحسب أعمارهم في انجلترا من هذا النوع إذ ان عدد المصابين يتزايد كلما كبرت الأعمار حتى يصل العدد اقصاه عند الكهول وذلك طبعاً لضعف حواسهم الطبيعية ويقظتهم عند هذا العمر المتأخر . أما في مصر فقد حاولنا الوقوف على طبيعة شكل المنحنى لتلك الظـــاهرة نفسها فوجدنا ان المنحنى التكراري شبيه بمثله في

انجلترا مع اختلاف بسيط هو تزايد التكرارات نوعاً ما عند الغمر ١٠ الى ١٥ سنة ، وهذا النوع من المنحنيات يطلق عليه اسم منحنى لامي .

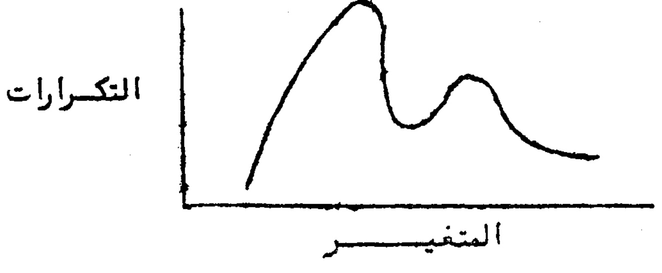
أما النوع الثاني من التوزيعات حاد الالتواء فهي ما كانت أكبر تكراراتها تقابل أصغر قيم المتغير ثم تبدأ التكرارات بعد ذلك في الانخفاض كلما زادت قيم المتغير حتىتصل التكرارات اقلها عند أكبر القيم كما هو مبين بالشكل التالي. ويطلق على هذا النوع من المنحنيات اسم رائي مقلوب Reversed J, shaped حيث انه عكس المنحنى الرائي السابق وصفه . أما إذا بدأت التكرارات تتزايد بعد وصولها الى نهايتها الدنيا فيطلق على هذا النوع اسم « منحنى لامي مقلوب ». وهذا النوع من المنحنيات كثير الحدوث في المسائل الاقتصادية الخاصة بتوزيع الثروة ، مثلاً توزيع الملاك الزراعين بحسب مساحة ما يملكونه من أراض من أحسن الأمثلة لهذا النوع . ومن أمثلة هذا التوزيع أيضاً توزيع الشركات المساهمة حسب رؤوس أموالها وكذلك التوزيع التكراري لحالات الطلاق التي تحدث سنويا في مصر موزعة حسب طول الحياة الزوجية التي سبقت الطلاق . ويكاد يكون التوزيع التكراري للسكان في مصر في الوقت الحاضر بحسب السن من هذا النوع ايضاً والشذوذ الوحيد في المثال الأخير هو أن عدد من هم دون الواحدة أقل نسبياً من عدد الفئة التي تليها مباشرة مما يجعل للمنحنى التكراري تقوس ضئيل إلىأسفل عند بدايته. ونلاحظ ان كثيراً ما نشاهد تلك الظاهرة (ظاهرة التقوس) في هذا النوع من المنحنيات كما نشاهد التقوس أيضاً في نهاية المنحنيات الرائية وكثيراً ما يعمد الاحصائيون إلى ضم الفئات الاولى من التوزيع إلى بعضها للتخلص من هذا التقوس .

ويسمى هذا المنحنى احياناً بمنحنى « باريتو » نسبة إلى وليفريد باريتو الذي وجه انتباه الاقتصاديين إلى هذا النوع في الدراسات الخاصة بتوزيع الثروة .



المنحنيات المتعددة القمم (Multi — Modal curves) :

وهي التي تتركز فيها القيم أي تعلو التكرارات عما حولها عند أكثر من نقطة واحدة كما هو واضح من الشكل الآتي :

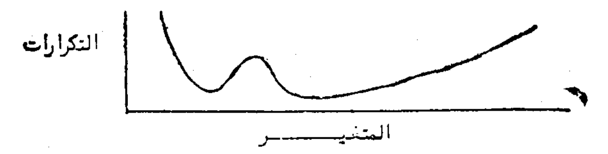


فإذا وجد توزيع تكراري من هذا النوع فان ذلك دليل قاطع على عدم تجانس افراد المجموعة المكونة له . فلو فرض أن أخذنا مجموعة كبيرة من الأشخاص الذين ينتمون إلى أجناس مختلفة لم قسمنا أطوالهم وكونا توزيعاً تكرارياً من هذه الأطوال فانه مما لا شك فيـــه ان التوزيع الناتج وبالتالي المنحني التكراري الذي يمثله يكون متعدد القمم حيث يتركز الطول لكل مجموعة عند نقطة معينة هي طول الغالبية العظمى لافراد كل جنس على حده. فلو فرض ان اقتصرنا في التوزيع على أطوال جنس واحد فـــان التوزيع التكراري الناتج يفقد خاصية تعدد القمم ويصبح توزيعاً وحيد القمة ، فتعدد

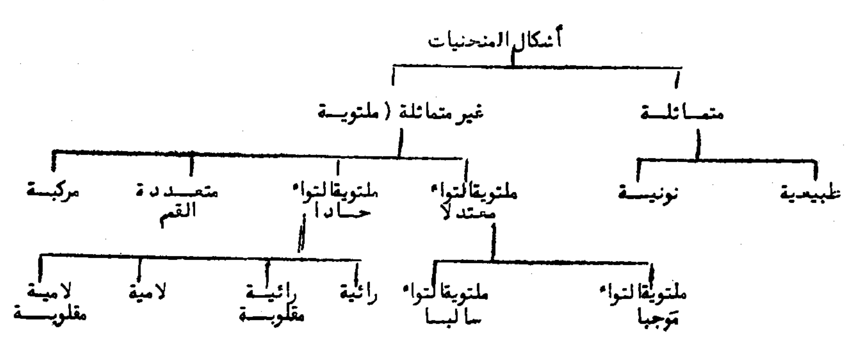
القمم في التوزيعات التكرارية اذن دليل مادي لا يخطىء على عدم تجانس المجموعة المكونة له . وكثيراً ما تستخدم هذه الحقيقة الهامة في البحوث العلمية للتحقق من تجانس المواد المختبرة أو عدم تجانسها .

فلو طلب منك ان تختبر نقاوة شحنة معينة من القمح مثلاً فيكفي أن تأخذ كمية محدودة من هذا القمح كعينة عشوائية ممثلة له وتكوّن توزيعاً تكرارياً لأوزان حبات القمح في تلك العينة ، فاذا كان المنحنى الذي يمثل أوزان هذه المجموعة منحنى متعدد القمم كان ذلك دليلاً قاطعاً على أن القمح الذي تختبره مخلوط بنوع آخر من القمح وعدد القمم ينم على عدد الاصناف المخلوطة . أما إذا كان القمج نقياً ومن صنف واحد فمن المحتمل أن يكون التوزيع الناتج متماثلاً والمنحنى ذا قمة واحدة فقط .

وثمة نوع آخر من المنحنيات الملتوية وهو ما يطلق عليه المنحنيات المركبة (Complex Curves) وهو ما تكونت أجزاؤه من عدة أنواع مختلفة من المنحنيات التي ذكرناها سابقاً . فلو أخذنا التوزيع التكراري للمتوفين في انكلترا حسب أعمارهم عند الوفاة لوجدنا التوزيع كما هو مبيّن في الشكل الآتي :



ويمكن تلخيص التقسيم السابق في الشكل الآتي :



تمثيل التوزيعات التكرارية ذات الفئات الغير متساوية :

قلنا عند الكلام على اختيار فئات التوزيعات التكرارية بوجوب أخذ فئات متساوية المدى إذ أن في ذلك تسهيل لمقارنة تكرارات التوزيع وتبسيط للتمثيل البياني لها. ولكننا ذكرنا في معرض الكلام عن هذا الموضوع أن هناك اعتبارات عملية لا تسمح باتباع تلك النصيحة حيث أن طبيعة التوزيع تتطلب أحياناً التفصيل في بعض أجزائه والاختصار في البعض الآخر، الأمر الذي يقتضي جعل الفئات غير متساوية .

يترتب على ذلك أن هناك بعض القواعد التي يجب أن نتبعها عند مقارنة التوزيعات التكرارية ذات الفئات الغير متساوية وتمثيلها بيانياً . إذ أنه لا بد من تعديل تكرارات الفئات المختلفة تبعاً لمدى كل منها إذ لا يجوز بحال من الأحوال مقارنة تكرارات فئتين أو أكثر مداها مختلف دون تعديل .

والمثل الآتي يوضح ضرورة تعديل تكرارات الفئات الغير متساوية قبل مقارنتها .

لو قيل مثلاً ان عدد العمال الذين يتقاضون اجراً يومياً قدره يتراوح بين ٢٠ قرشاً و ٣٠ قرشاً هو ٥٠ عاملا وان عدد ما يتقاضون أجراً بين ٣٠ ، ٦٠ قرشاً هو ١٢٠ عاملاً وسئلت بعد ذلك أي الفئتين أكبر عدداً — عمال الأجور الصغيرة أو عمال الأجور الكبيرة — فماذا يكون جوابك ؟ هل يصح أن تقول بما أن عمال الأجر العالي ١٢٠ عاملاً أكثر من عمال الأجر المنخفض ٥٠عاملاً فان أصحاب الأجور المنخفضة أقل عدداً من أصحاب الأجور المرتفعة. الواقع أن عدد أصحاب الأجر المنخفض قليلون لا لسبب الا لأن مدى فئتهم صغير (مدي ١٠ قروش فقط ، أي من ٢٠ إلى ٣٠ قرشاً) في حين أن مدى الأجور العالية هو ٣٠ قرش (أي من ٣٠ الى ٦٠ قرشاً) . فلو فرض ان توزيع أصحاب الأجور العالية في فئتهم منتظم فمن الواضح ان عدد من يتقاضى أجراً من ٣٠ الى ٤٠ قرشاً يساوي عدد من يتقاضى اجراً من ٤٠ إلى ٥٠ قرشاً يساوي عدد من يتقاضى اجراً من ٥٠ الى ٦٠ قرشاً ، والعدد في كل

حالة هو ٤٠ عامل ومن ذلك نستنتج أن أصحاب الأجور المنخفضة في هذا المصنع أكثر عدداً من أصحاب الأجور العالية وهو عكس مـا وصلنا اليه بمجرد مقارنة التكرارات الأصلية دون الاهتمام بمدى فئاتهم .

يستنتج من هذا انه لا بد من تعديل تكرارات ألفئات الغير متساوية قبل مقارنة تلك التكرارات.وذلك بأن نأخذ طول فئة معينة كوحدة ثم نضرب تكرارات كل فئة في نسبة مدى الفئة المتخذة وحدة الى مدى تلك الفئة . فلو أخذنا مدى الفئة ١٠ في المثال السابق كوحدة ثم ضربنا تكرارات الفئة الثانية في النسبة بين مدى الوحـــدة ومدى تلك الفئة نحصل بذلك على التكرار المعدل .

$$\text{التكرار الأصلي} \times \frac{\text{طول الفئة المتخذة كوحدة}}{\text{طول الفئة المراد تعديلها}} = \text{التكرار المعدل}$$

$$١٢٠ \times \frac{١٠}{٣٠} = ٤٠$$

وإظهاراً لأهمية تلك القاعدة وتطبيقاً لها نعود إلى كشف البيانات الأصلية الخاص بأجور العمال ونكوّن منه توزيعاً تكرارياً تكون فئاته غير متساوية المدى ثم نمثل التوزيع بطريقتين الأولى خاطئة أن دون تعديل في التكرارات والثانية صحيحة وهي التي تم فيها هذا التعديل ، وسوف نرى اننا سنحصل على توزيع يختلف قانون تغيره اختلافاً تاماً عن قانون تغير التوزيع كما نعرفه. وسيرى القارىء أن تعديل التكرارات وفقاً للقاعدة التي ذكرناها سيعيد قانون التغير الى حقيقته .

فئات الأجور

٤ — ٦ — ٨ — ١٠ — ١٢ — ١٤ — ١٦ — ١٨ — ٢٠ — ٢٤ — ٣٠ — ٥٢

التكرار (عدد العمال)

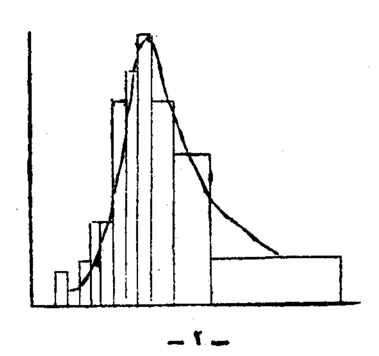
٤ ٢ ٥ ٩ ٩ ٢٢ ٢٥ ٢٩ ٤٣ ٤٨ ٥٤

يذكر القارىء أن فائدة عمل توزيع تكراري تنحصر في ثلاث نقاط :

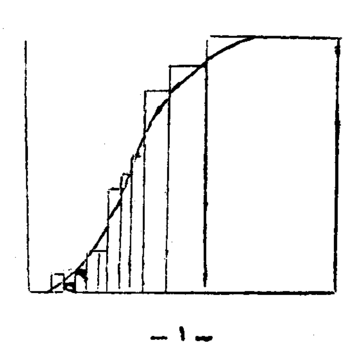
الأولى تحديد مدى التغير ، والثانية الكشف عن وجود التركز في القيم والقيمة التي يحدث عندها التركز ، والثالثة الوقوف على الاستمرار في التغير بين حدى المدى .

وفضلاً عما تقدم فان الرسم البياني على اختلاف انواعه سواء كان أشرطة أو مدرجات أو مضلعات او منحنيات يعطى للعين صورة مجسمة لقانون التغير في التوزيع موضوع الدراسة. ونعني بقانون التغير طريقته ، أي العلاقة بين التكرارات والتغير .

فإذا أخذنا توزيع الأجور السابق مثلاً نرى انه كلما زاد الاجر زاد معه تكراره في بادىء الأمر حتى يصل التكرار الى أقصى حدوده عند حوالي ١٨ ليرة ثم يبدأ التكرار يتناقص كلما زاد الأجر ولكن بسرعة أقل بكثير من السرعة التي ارتفع بها حتى وصل الى قمته . ومعنى ذلك ان الأجر في هذا المصنع « يميل » نحو الاجور الصغيرة ، اذ أن معظم التكرارات الكبيرة تلتف حول القيم الصغيرة للأجور .



— ٢ —



— ١ —

واضح من الرسم الاول (١) ان شكل المنحنى رائي وهـــذا يختلف اختلافاً بيناً عن الشكل الذي سبق أن حصلنا عليه للتوزيع نفسه عند تقسيمه الى فئات متساوية .

وغني عن القول ان التغيير في اختيار مدى الفئات لا يجب أن يؤثر على قانون تغير المجموعة او شكل منحناها . فذلك ثابت مهما اختلفت الفئات ، والسبب الحقيقي للنتيجة الخاطئة التي وصلنا اليها هو عدم تعديل التكرارات وفقاً للقاعدة التي شرحناها .

أما الجدول الآتي فيعطى التوزيع وتكراراته المعدلة . فاذا قمنا بتعديل التكرارات كما هو مبين في الجدول ورسمنا المدرج والمنحنى التكراري للتوزيع كما هو واضح بالشكل رقم (٢) فاننا نحصل على منحنى ملتو التواء موجباً وهو الشكل الحقيقي للتوزيع كما نعرفه .

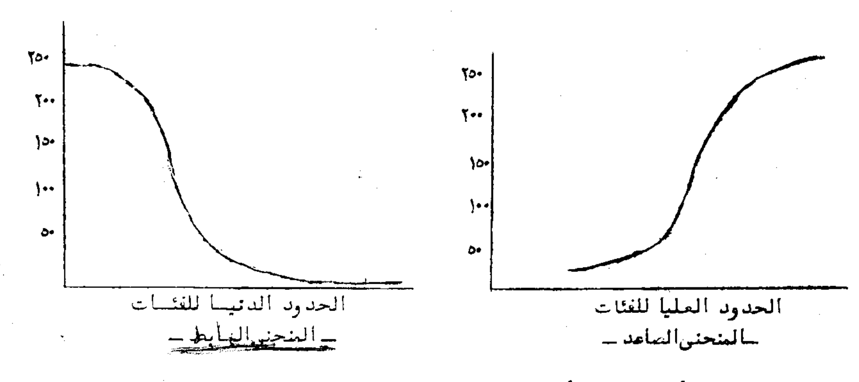
فئات الأجور	التكرارات الأصلية (عدد العمال)	أطوال الفئات	معامل التصحيح = طول الوحدة / طول الفئة	التكرارات المعدلة = التكرار الاصلي × معامل التصحيح
٤ —	٤	٢	١	٤
٦ —	٢	٢	١	٢
٨ —	٥	٢	١	٥
١٠ —	٩	٢	١	٩
١٢ —	٩	٢	١	٩
١٤ —	٢٢	٢	١	٢٢
١٦ —	٢٥	٢	١	٢٥
١٨ —	٢٩	٢	١	٢٩
٢٠ —	٤٣	٤	٢ / ٤	٢١ر٥
٢٤ —	٤٨	٦	٢ / ٦	١٦
٣٠ — ٥٢	٥٤	٢٢	٢ / ٢٢	٤ر٩
المجموع	٢٥٠	—	—	—

المنحنيات التكرارية المتجمعة :

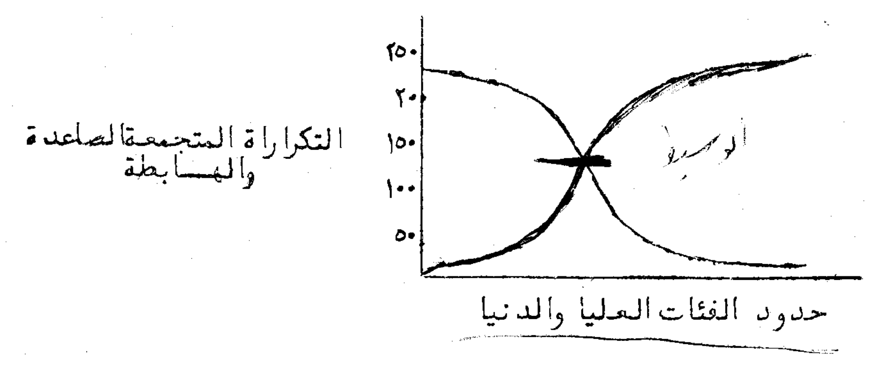
ذكرنا سابقاً اننا نستفيد أحياناً بتجميع التكرارات اما صعوداً أو هبوطاً فنحصل على التوزيع التكراري المتجمع الصاعد أو الهابط . وتظهر

فائدة هذا النوع من التوزيعات بشكل واضح اذا مثلناها بيانياً كما فعلنا في التوزيعات التكرارية البسيطة .

وطريقة توضيح هذه التوزيعات تتلخص في تمثيل الحدود العليا للفئات على المحور الأفقي (المحور السيني) والتكرارات المتجمعة الصاعـدة على المحور الرأسي (المحور الصادي) ثم نضع نقطاً تمثل العلاقة بين حدود الفئات العليا وتكراراتها ثم نوصل النقط الناتجة بخط ممهد فنحصل على الشكل الآتي . ويطلق على المنحنى في هذه الحـالة اسم المنحنى التكراري المتجمع الصاعد ويلاحظ بنوع خاص أن تكرارات كل فئة لا تقابل مراكز الفئات كما كان الحال في المضلع التكراري بل انها تقابل الحدود العليا للفئات . وبالمثل يمكن تمثيل المنحنى المنحنى المتجمع الهابط بيانياً بتخصيص المحور الأفقي للحدود الدنيا للفئات والمحور الرأسي للتكرارات المتجمعة الهابطة ثم تحديد النقط الدالة على العلاقة بين الحدود الدنيا والتكرارات من واقع الجدول ثم توصيل تلك النقط بخط ممهد كما هومبين بالشكل ويطلقعلى هذا المنحنى اسم المنحنى التكراري المتجمع الهابط، ويجب الأشارة هنا كذلك الى أن تكرارات كل فئة تقابل الحدود الدنيا للفئات وليست مراكزها كما كان الحال في الرسوم البيانية التكرارية البسيطة .



ولا يفوتنا أن نذكر أنه من الممكن تمثيل المنحنيان المتجمعان الصاعد والهابط في شكل واحد كما هو مبين بالرسم التالي . ويلاحظ أن نقطة تقاطع هذين المنحنيين ذات أهمية خاصة اذ يقطع العمود النازل منها المحور الأفقي في نقطة تمثل أحد متوسطات الأجر، اذ أنها النقطة الوحيدة التي يتساوى عندها عدد الذين يتقاضون أجراً أقل منها وعدد الذين يتقاضون أجراً أكبر منها ، وسنطلق على هذا النوع من المتوسطات اسم « الوسيط » (Median) كما سنشرح ذلك عند الكلام على موضوع المتوسطات .



وفائدة هذه المنحنيات هي نفس فائدة التوزيعات المتجمعة فيمكن منها معرفة عدد العمال الذين يتقاضون أقل من أجر معين أو أكثر من أجر معين ، كما يمكن معرفة الأجر الذي يتقاضى دونه أو أكثر منه نصف العمال أو ربعهم أو أي عدد معين من العمال .

مقارنة التوزيعات التكرارية :

لو طلب منا مقارنة التوزيع التكراري لأجور عمال مصنعين كالآتي :

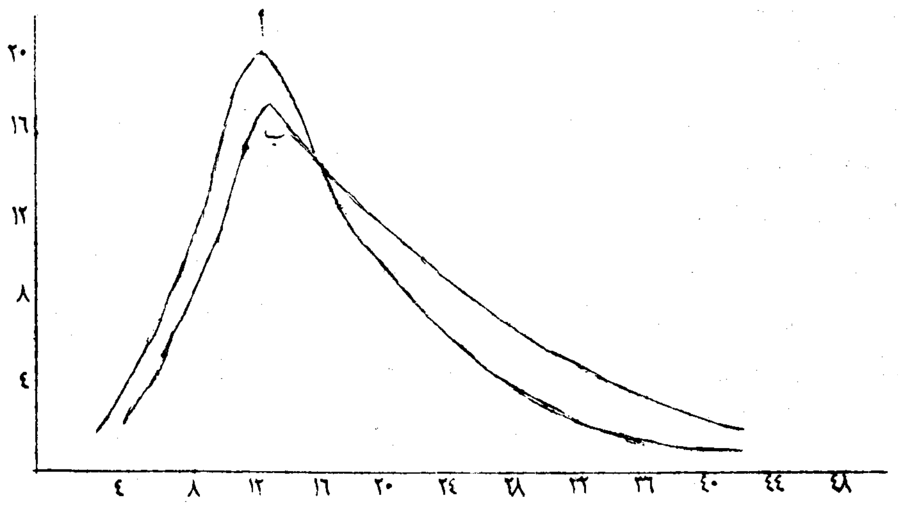
فئـات الأجور	عمـال مصنع أ	عمال مصنع ب
٤ —	٦	٧
٨ —	١٤	٢٥
١٢ —	٣١	٦٨
١٦ —	٥٤	١١٦
٢٠ —	٤٣	٩٦
٢٤ —	٣٦	٨٤
٢٨ —	٢٧	٦٨
٣٢ —	١٩	٥٦
٣٦ —	١٠	٤٠
٤٠ —	٦	٢٤
٣٤ —	٣	١٣
٤٨ — ٥٢	١	٣
	٢٥٠	٦٠٠

فخير الوسائل المتبعة لمقارنة مثل هذين التوزيعين هو تمثيلها في شكل بياني . ويجب استخدام المنحنى التكراري دون المدرج التكراري إذ يمكن رسم المنحنيين على نفس الشكل .

واضح من الشكل السابق أن مدى التغير في الأجور واحد تقريباً في المصنعين وان التغير مستمر في الحالتين بين حدى المدى وان نقطة التركز تكاد تكون واحدة حيث انها تتراوح بين ١٤ ، ١٨ ليرة في المصنعين . ولكن بالنظر إلى الشكل لا يمكننا مقارنة نسبة عدد العمال الذين يتقاضون أجوراً معينة اذ أن الشكل يحملنا على الاعتقاد لأول وهلة ان نسبة العمال الذين يتقاضون فئات الأجور المتوسطة أكبر بين عمـــال مصنع ب منها بين عمال مصنع أ ، ولكن الواقع غير ذلك إذ أن نسبة العمال الذين يتقاضون الأجور الصغيرة والمتوسطة بين عمال مصنع أ أكبر منها بين عمال مصنع ب وعلى العكس نسبة العمال الذين يتقاضون أجوراً عالية في مصنع أ أقل منها في مصنع ب . ولا يمكننـــا الوقوف على تلك الحقائق إلا بعد تحويل أرقام التوزيعين السالفين إلى نسب مئوية حيث ان مجموع عدد العمال في العينتين غير متساو . والجدول التالي يعطي التوزيعين بتكرارتها العادية والمئوية بالنسبة للمجموع في كل مصنع .

فئات الأجور	عمال مصنع أ	النسب المئوية لهم	عمال مصنع ب	النسب المئوية لهم
٤ —	٦	٢,٤	٧	١,١٦
٨ —	١٤	٥,٦	٢٥	٤,١٦
١٢ —	٣١	١٢,٤	٦٨	١١,٣
١٦ —	٥٤	٢١,٦	١١٦	١٩,٣
٢٠ —	٤٣	١٧,٢	٩٦	١٦,٠
٢٤ —	٣٦	١٤,٤	٨٤	١٤,٠
٢٨ —	٣٧	١٠,٨	٦٨	١١,٣
٣٢ —	١٩	٧,٦	٥٦	٩,٣
٣٦ —	١٠	٤,٠	٤٠	٦,٦
٤٠ —	٦	٢,٤	٢٤	٤,٠
٤٤ —	٣	١,٢	١٣	٢,١٦
٤٨ — ٥٢	١	٠,٤	٣	٠,٤٩
	٢٥٠	١٠٠	٦٠٠	١٠٠

فاذا أوضحنا التوزيـع التكراري للنسب المئوية في كل من المصنعين نحصل على الرسم الآتي :



ويلاحظ القارىءأن جميع فئات التوزيعيين التكراريين السابقين متساوية، فاذا فرض أن كان لدينا توزيعين فئاتها غير متساوية المدى فانه من الواجب تعديل التكرارات الأصلية وبالتالي النسب المئوية لتلك التكرارات تبعاً لمدى كل فئة بالطريقة المعتادة قبل عمل الرسم البياني .

كذلك يلاحظ القارىء أن المقارنة تمت بين توزيعين طبيعة مميز المتغيرين في كل منهما واحدة (الليرة مثلاً) . ولكن لا يمكن عمل مقارنة من هـذا النوع بين توزيعين يختلف مميز المتغير فيهما .

وقد يطلب منا في بعض الأحيان مقارنة توزيعين تكراريين متجمعين ، وفي هذه الحالة لا يجب الاعتماد على المنحنيين المتجمعين بـــل يلزم تحويل التكرارات المتجمعة لكل من التوزيعيين الى نسب مئوية ورسم المنحنيين من واقع هذه النسب . ويلاحظ عند مقارنة التوزيعات التكرارية المتجمعة انه لا تأثير لاختلاف مدى الفئات على التكرارات المقارنة بعكس الحـال في مقارنة التوزيعـــات التكرارية البسيطة . أما فيما يختص بمقارنة التوزيعات التكرارية المتجمعة ذات المتغيرات المختلفة فانه يستحيل علينا مقارنة منحنياتها المتجمعة أسوة بمنحنياتها العادية .

منحنى لورنز :

يمكننا الاستفادة من عمل التوزيع التكراري لأي متغير بجانب ما ذكرنا في الوقوف على درجة « التفاوت » بين أفراد المجموعة وذلك بتمثيل التوزيع بمنحنى جديد يطلق عليه اسم منحنى « لورنز » (Lorenz Curve) اشادة باسم مبتكره .

لنفرض اننا نريد عمل منحنى لورنز لتوزيع الأجور المعروف فينحصر العمل في الخطوات التالية :

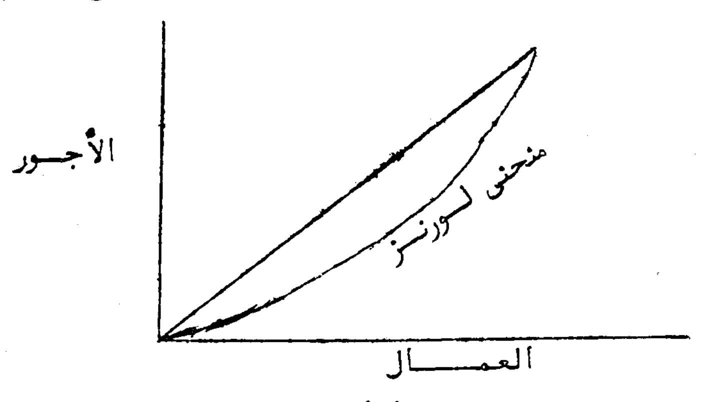
١ – ايجاد الأجور التي يتقاضاها أفراد كل فئة على حدة بضرب عدد عمال الفئة في مركزها .

٢ – عمل توزيعين متجمعين أحدهما للتكرارات الأصلية (أي العمال) والآخر للاجور التي يتقاضاً هؤلاء العمال (حاصل ضرب تكرار كل فئة × مركزها) .

٣ – حساب النسب المئوية لكل فئة من فئات التوزيعين ثم رصد هذه النسب على المحورين الرأسي والأفقي فيكون الخط الذي يمثل العلاقة بينهما هو منحنى لورنز المطلوب .

فئات الأجور	التكرارات	المراكز	التكرارات × المراكز	التكرارات متجمعة	الأجور متجمعة	تكرارات مئوية متجمعة	أجور مئوية متجمعة
٤ –	٦	٦	٣٦	٦	٣٦	٢,٤	٠,٦
٨ –	١٤	١٠	١٤٠	٢٠	١٧٦	٨,٠	٣,١
١٢ –	٣١	١٤	٤٣٤	٥١	٦١٠	٢٠,٤	١٠,٦
١٦ –	٥٤	١٨	٩٧٢	١٠٥	١٥٨٢	٤٢,٠	٢٧,٦
٢٠ –	٤٣	٢٢	٩٤٦	١٤٨	٢٥٢٨	٥٩,٢	٤٤,٠
٢٤ –	٣٦	٢٦	٩٣٦	١٨٤	٣٤٦٤	٧٣,٦	٦٠,٣
٢٨ –	٢٧	٣٠	٨١٠	٢١١	٤٢٧٤	٨٤,٤	٧٤,٥
٣٢ –	١٩	٣٤	٦٤٦	٢٣٠	٤٩٢٠	٩٢,٠	٨٥,٧
٣٦ –	١٠	٣٨	٣٨٠	٢٤٠	٥٣٠٠	٩٦,٠	٩٢,٣
٤٠ –	٦	٤٢	٢٥٢	٢٤٦	٥٥٥٢	٩٨,٤	٩٦,٧
٤٤ –	٣	٤٦	١٣٨	٢٤٩	٥٦٩٠	٩٩,٠	٩٩,١
٤٨ – ٥٢	١	٥٠	٥٠	٢٥٠	٥٧٤٠	١٠٠,٠	١٠٠,٠
المجموع	٢٥٠	–	٥٧٤٠	–	–	–	–

لو لم يكن هناك تفاوتاً بين نسبة العمال ونسبة الاجورالتي يتقاضونها بمعنى أن ٢٠ ٪ من العمال يتقاضون ٢٠٪ من الأجور وان ٤٠٪ من العمال يتقاضون ٤٠٪ من الأجور وهكذا، لانطبق منحنى لورنز على قطر المربع المرسوم فيه كما



هو واضح في الرسم ، ولكن نظراً لأن الحقيقة غير ذلك إذ ان ٢٠ر٤ من العمال يتقاضون ١٠ر٦ ٪ من مجموع الاجور وان ٤٢ ٪ من العمال يتقاضون ٢٧ر٦ ٪ من الاجور فإن المنحنى تقوس الى أدنى بعض الشيء ويدل مقدار التقوس على مدى التفاوت بين التوزيعين - العمال وأجورهم - ويمكن تظليل المساحة بين الخط المستقيم والمنحنى لإظهار مدي التفاوت في التوزيعين فكلما زادت هذه المساحة المظللة كلما دل ذلك على تفاوت أكبر .

ويستخدم هذا المنحنى بكثرة في التوزيعات التكرارية التي يهم الباحث الوقوف على درجة التفاوت فيها . فيهمنا في بعض الحالات مقارنة عدد المصانع وكمية ما تنتجه لاظهار تركز الانتاج في وحدات قليلة أو انعدام هذا التركز وكذلك توزيعات الدخل بين أفراد المجتمع لاظهار تركز الثروة في يد أفراد قلائل أو انعدام هذا التركز .

وكذلك يمكن عمل منحنيين من هذا النوع ورسمها في شكل واحد لمقارنة مدى التفاوت في التوزيعين اللذين نقوم ببحثهما . (قارن التفاوت في الأجور في المصنعين أ ، ب) . ونلاحظ أنه إذا تعادل التفاوت في التوزيعين فيستحسن رسم المنحنيين على جانبي خط التماثل وذلك باستخدام المحور الأفقي

للعمال في المصنـــع أ والمحور الرأسي لأجورهم ، ثم استخدام المحور الأفقي للأجور والرأس للعمال في المصنع ب ، وذلك لامكان عمل المقارنـة بينهما ، أما إذا كانت نسبة التفاوت تختلف كثيراً بين التوزيعين يمكن رسم المنحنيين على جانب واحد من خط التماثل حيث تسهل المقارنة في هذه الحالة .

التصوير الجبري للخطوط البيانية :

سبق أن ذكرنا ان الخط الذي يتوسط النقط التي نرصدها في الرسم البياني يسمى خط الاتجاه العام للظاهرة موضوع البحث ، حيث انه يدل على اتجاه التغير في قيم هذه الظاهرة في الفترات الزمنية المختلفة . كذلك إذا كان الرسم يوضح ظاهرتين سوياً فان خط الاتجاه العام يكون هو الخط الذي يمثل اتجاه التغير في احدي الظاهرتين تبعاً للتغير في الظاهرة الاخرى . ومن الواضح أن نوع التغير في الظاهرة سواء كان تغيراً زمنياً أو تغيراً ناتجاً عن تغير ظاهرة اخرى لا بد أن يؤثر على شكل الخط الذي يمثل الاتجاه العام .

وخط الاتجاه العام يمكن أن يكون خطاً مستقيماً اما صاعداً نحو اليمين أو هابطاً نحو اليمين ويدل على أن نوع التعبر في الظاهرة موضوع البحث ثابتاً دائماً فهي اما في ازدياد مستمر إذا كان صاعداً نحو اليمين أو نقص مستمر اذا كان هابطاً نحو اليمين . ان هذا الشكل المنتظم لخط الاتجاه العام جعل الرياضيين يحاولون وضع معادلة تدل عليه وقد توصلوا الى ان معادلة من الدرجة الاولى تعبر عن اتجاهه ، وتتخذ هذه المعادلة الصيغة الآتية: ص = م س + ب حيث م تدل على ميل الخط وتكون موجبة إذا كان الخط صاعداً نحو اليمين وسالبة إذا كان الاتجاه العام هابطاً نحو اليمين ، أما ب فقيمة ثابتة تدل على الجزء الذي يقطعه خط الاتجاه العام من المحور الرأسي .

وعندما يكون الخط معبراً عن الاتجاه العام للظاهرة في فترات زمنية مختلفة تكون س في المعادلة رمزاً لهذه الفترات، أما إذا كان الخط معبراً عن

الاتجاه العام للعلاقة بين ظاهرتين تكون س ، ص رموزا لهما وفي هذه الحالة يمكن أن يوجد خطان للاتجاه العام ، احدهما يصور اتجاه الظاهرة ص في علاقتها مع الظاهرة الاخرى س ، ويظهر في المعادلة ص = مَ س + ب ، والثاني يصور اتجاه الظاهرة س في علاقتها مع الظاهرة ص ، ويظهر في المعادلة س = مَ ص + بَ وفي هذه المعادلة تكون مَ هي كذلك رمزاً لميل خط اتجاه س بالنسبة للظاهره ص ، بَ رمزا للجزء المقطوع من المحور س بواسطة خط الاتجاه العام .

وقد سبق أن ذكرنا أن خط الاتجاه العام يجب أن يتوسط النقط التي تمثل الاتجـــاه العام للظاهرة في تغيرها الزمني أو الاتجاه العام للعلاقة بين ظاهرتين ، ومن الواضح أنه من الصعب رسم مثل هذا الخط بحيث يتوفر فيه شرط التوسط، الا انه على أساس معادلة الاتجاه العام يمكن تحديد النقط التي تقع عليه وبذلك يمكن تحديد اتجاهه الوسطي بين النقط الواقعية بدقة تامة . والمشكلة هي في كيفية ايجاد هذه المعادلة ويستخدم لذلك ما نسميه بطريقة المربعات الصغرى حيث أننا بهذه الطريقة نوجد معادلة أفضل خط يمثل الاتجاه العام وهو الخط الذي يتوفر فيه الشرطان :

١ – ان مجموع انحرافات النقط الواقعية عن هذا الخط = صفر .

٢ – ان مجموع مربع انحرافات النقط الواقعية عند هذا الخط أصغر من مجموع مربعات انحرافات النقط الواقعية عن أي خط آخر .

وسوف نعود الى مناقشة هذه الطريقة في موضوع الانحدار .

كذلك يمكن أن يكون الاتجاه العام متخذاً شكل منحني من الدرجة الثانيـــة وتصوره المعادلة ص = أ س^{2} + ب‌س + حـ حيث أ ، ب ، حـ ثوابت معينة . وتمثل هذه المعادلة ما نسميه بالقطع المكافىء ، وهو اما محدباً الى أعلى أو مقعراً الى أعلى ، فاذا كانت قيمة أ سالبة يكون القطع المكافيء

محدباً الى أعلى ، اما اذا كانت موجبة يكون القطع مقعراً الى أعلى . وسواء كانت ا موجبة أو سالبة فان النقطة التي يصل القطـع المكافيء الى قيمته الكبرى أو الصغرى عندها هي النقط التي تتحدد على أساس المعادلين :

$س = - \frac{ب}{٢ أ}$ ، $ص = \frac{٤ أ حـ - ب^{٢}}{٤ أ}$. هذه النقطة تسمى رأس القطع المكافيء. ومن الواضح أن الشكل العام لمعادلة الدرجة الثانية يمكن أن يختلف تبعاً لما اذا كان القطع المكافيء يقطع محور س مرتين واما أن يقطعه مرة واحدة أي يمسه، واما أن لا يقطعه بتاتاً فيقع بتمامه فوق محور س أو تحته .

ومنحنى القطع المكافيء يجب كذلك أن يتوسط النقط التي تمثل الاتجاه العام ولهذا يمكن استخدام طريقة المربعات الصغرى في ايجاد معادلته .

كذلك يمكن أن يكون الاتجاه العام من الدرجة الثالثة فتمثله المعادلة .

$ص = أ س^{٣} + ب س^{٢} + حـ س + د$

كذلك يمكن أن يكون الاتجاه العام من درجة أعلى .

ومن الواضح أن المنحنيات التكرارية يمكن التعبير عنها بمعادلات وسنترك ذلك في الوقت الحاضر حتى ندرس الخصائص الرياضية للتوزيـع التكراري .

القواعد العامة للرسم البياني :

الرسم البياني هو كما ذكرنا الوسيلة الثانية التي يلجأ اليها الاحصائي لتوضيح البيانات بعد جمعها وتبويبها حيث أن الأرقام المبوبة في جداول قد تكون صعبة الفهم والاستيعاب بالنسبة لكثير من الناس . لذلك يجب أن يكون الرسم البياني واضحـــاً من جميع النواحي حتى يمكن أن يؤدي وظيفته التوضيحية بالنسبة لمن يطلع عليه .

كذلك تساعد بعض أنواع من الرسومات البيانية في اجراء تحليلات احصائية معينة . وفي هذه الحالة يجب أن يتوفر في الرسم الدقة.(ليست الدقة الهندسية الكاملة) حيث ان استنتاجنا من الرسم عندما لا يكون دقيقاً يمكن أن يكون خاطئاً مضللا .

لذلك يكون الوضوح التام والدقة الممكنة هما الصفتان الأساسيتان في الرسوم البيانية الاحصائية . وحتى يمكن أن يتوفر في الرسم هاتان الصفتان يجب أن نلاحظ القواعد الآتية عند تصميمه وتنفيذه :

أولاً :

لكل نوع من البيانات الاحصائية المبوبة رسم خاص بها يكون قادراً على توضيحها وابراز الفكرة الكامنة وراءها ، حيث اننا بأي رسم بياني نريد أن نقول لمن يطلع عليه شيئاً معينـــاً . هذا الشيء هو الفكرة الكامنة وراء البيانات التي نقوم بتوضيحها.قد يكون غرضنا من الرسم أن نبرز طريقة التغير في الظواهر موضوع البحث خلال فترة زمنية معينة ، وقد يكون غرضنا أن نظهر تقسيم بيان كلي إلى أجزائه المختلفة في عام واحد ، وقد يكون غرضنا أن نظهر هـــذا التقسيم في عدة أعوام ليتضح التغيير في تركيب الظاهرة من عام الى آخر ، وقد يكوق غرضنا أن نظهر المقارنة في طريقة التغير بين ظواهر مختلفة خلال فترة زمنية معينة ، وقد يكون غرضنا أن نظهر مقدار ما حدث من زيادة أو نقصان في الظاهرة أو الظواهر موضوع البحث ، وقد يكون غرضنا أن نظهر معدل ما حدث من زيادة أو نقصان في هذه الظواهر . الـــخ . لذلك تكون أول خطوة في تصميم الرسم البياني الاحصائي هي تحديد نوع الرسم الذي يتفق مع البيانات التي تريد توضيحها ومع الفكرة التي نريد ابرازهــا . ان الهدف من الرسم لا يمكن أن يتحقق تماماً الا باختيار النوع المناسب .

ثانياً :

ان الفكرة الأساسية من الرسم البياني الاحصائي هي التوضيح لذلك يجب أن يتوفر فيه السهولة والذوق السليم من جميع النواحي . فإذا كان الرسم يحتوي على خطوط بيانية مثلاً يجب أن تظهر ثقيلة واضحة بحيث تبرز للعين دون حاجة الى فحص وتدقيق (وذلك بعكس الرسوم البيانية الهندسية حيث تكون الخطوط فيها رقيقة خفيفة) . كذلك إذا احتوى الرسم على عدة خطوط بيانية يجب التميز بينها بالألوان أو باي طريقة اخرى وفي هذه الحالة يجب أن يصحب الرسم توضيح لمعاني هذه الألوان أو الطرق المختلفة التي استعلمت . كذلك إذا أردنا أن نستعمل الألوان لا يجب أن تفعل ذلك الا إذا كان لدينا خبرة تامة باستعمالها . وبشكل عام كل ما يساعد في جعل الرسم مشوقاً يجب استعماله .

ثالثاً :

إن القيم التي نريد توضيحها بالرسومات البيانية الاحصائية قد لا يمكن أن نظهرها بحيث يتوفر فيها الدقة الحسابية التامة ولذلك نضطر أحياناً الى تقريب الأرقام حتى نظهرها ، إذ المقصود من أي رسم احصائي هو اعطاء فكرة توضيحية لمن يطلع عليه ، وليس اظهار القيم بدقة حسابية متناهية فإذا كان لدينا القيمة ٥٨٣٧٤٥٤٣ مثلاً نلاحظ أنه ليس من المعقول أن نحاول اظهار هذه القيمة بكاملها على الرسم البياني العادي وبذلك نضطر الى تقريبها الى أقرب مائة ألف حتى نستطيع أن نحدد موقع النقط التي تعبر عنها تبعاً لمقياس الرسم الذي نتخذه .

رابعاً :

يجب البدء بالصفر على المحور الرأسي إلا إذا كان الرسم نصف لوغاريتمي ، وكذلك يجب أن يستمر مقياس الرسم منتظماً ، أما إذا تغير

مقياس الرسم يجب الاشارة الى ذلك في مكانه باستخدام عـلامة مميزة // . وبشكل عام لا يجب أن نغير من مقياس الرسم على المحور الرأسي إذا كان جزء من الرسم يدخل ضمن المقياس المختلف ، وبمعنى آخر إذا اضطررنا الى ذلك تبعاً لطبيعة الأرقـام يجب أن نلاحظ البدء بالصفر ثم نضع العلامة // بحيث تكون أقل قيمة نريد توضيحها لاتقل بأي حال عن القيمة التالية التي نضعها بعد هذه العلامة (//) . أما المحور الأفقي حيث انه غالباً يعبر عن الزمن فيمكن البـدء بأي سنة نريدها بحيث يتسع فراغ الرسم لجميع السنوات المطلوب عرض البيانات الخاصة بها .

خامساً :

يجب أن تبرز النقط التي تحدد اتجاه الخط البياني بحيث تكون أكثر وضوحاً من الخط نفسه .

سادساً :

يجب أن نتحاشى توضيح بيـانات كثيرة ومتنوعة في رسم بياني واحد لان ذلك يؤدي الى التعقيد وعدم الوضوح وبذلك يفقد الرسم ميزته .

سابعاً :

إذا كان الرسم خاصاً بفترة زمنية تبدأ بسنوات وتنتهي باشهر السنة الأخيرة يجب وضع فاصل في نهاية البيانات السنوية يكون موازياً للمحور الرأسي وبذلك يكون الرسم في الوقع مكون من جزأين ، الجزء الأول ويظهر البيانات السنوية والجزء الثاني ويظهر البيانات الشهرية . وتستخدم هذه الطريقة بشكل عام إذا اختلفت الوحدات الزمنية في السلسلة التي نريد توضيحها.

ثامناً :

يجب توضيح عنوان الرسم توضيحاً يظهر الموضوع والتاريخ والمكان

المتعلق بالرسم ، وبشكل عام يكون عنوان الرسم هو نفس عنوان الجدول الذي يحتوي على البيانات التي نقوم بتوضيحها .

تاسعاً :

يجب الاشارة في أسفل الرسم الى مصدر البيانات التي يعتمد عليها . ويجب ذكر المصدر متضمناً اسم الكتاب واسم المؤلف ورقم الصفحة وتاريخ النشر ، وكذلك الحال إذا كان المصدر نشرة احصائية .

عاشراً :

يجب توضيح مميز كل من المحورين – الأفقي والرأسي – بحيث يظهر على كل منهما دلالته الخـاصة به

حادي عشر :

يجب توضيح مقياس الرسم في مكان واضح من الشكل البياني :

تمـــارين

١ – تريد إحدى المؤسسات الاحصائية القيام بدراسة عن معيشة الطالب الجامعي في بيروت . اكتب تقريراً توضح فيه الخطوات التي يجري العمل تبعاً لها ، مع تصميم الاستمارة التي ترى استخدامها لجمع المعلومات .

٢ – ناقش الأخطاء المختلفة التي تتعرض لهـا الدراسات الاحصائية وبين كيف يمكن التغلب عليها .

٣ – اكتب مـا تعرفه عما يأتي :

المجتمع الاحصائي والوحدة الاحصائية –، مجتمع المعاينة ووحدة المعاينة– دراسة احصائية وصفية بالعد الشامل – تضليل البيانات الاحصائية –الأساس

القانوني للدراسات الاحصـــائية ــ الرسم البياني كوسيلة لتوضيح البيانات الاحصائية .

٤ – في دراسة عن ميزانية الاسرة اشتملت استمارة البحث على سؤال عن قيمة الانفاق على الطعام وقد جاءت الاجابات كالآتي (بالدينار) .

١٠	٣٥	٨	٢٥	١٣	١٨	٤٤	٥٣	٤٥	١٢	٣٢
٢٧	٢٨	٤٠	١٤	٩	٥٢	١١	٤٣	٣٦	٢٦	١٦
٥	١٧	٥٤	٤١	١٤	١٥	٣٦	٣٤	١٠	٢٢	٢٥
٢٧	٢١	٢٣	١٢	٢٤	٢٠	٢٧	٢٨	٢٩	٧	٢٠
٢٣	٢٧	٣١	٢١	٤١	٢٧	٢٢	٢٥	٣٠	٥١	٤٢
٣٧	٢٣	١٥	٢١	٢٦	٢٥	٢٠	٤٤	٥١	٤١	١٧
٢٥	٢٨	٢٢	١٦	٢١	٤٨	٢٣	٢١	٢٠	٢٦	٢٩
٢١	٢٨	٢٧	٢٧	٢٦	٩	٢١	٣٠	١٨	٤١	٥٠
٢٢	٢٤	٢٩	٢٨	٣٠	٢٧	٢٥	٢٣	٥٠	٢٧	٢٦
٢٩										

المطلوب :

تبويب هذه البيانات في توزيـع تكراري تبدأ فئاته بالقيمة ٥ ومدى الفئة فيه ٥ .

٥ – يجب قبل البدء في أي تعداد اتخاذ بعض القرارات الأساسية . ما هي هذه القرارات ؟ ولماذا يجب اتخاذها أولاً ؟ وضح اجابتك بأمثلة .

٦ – الآتي توزيع السكان تبعاً لفئات العمر في احدى الدول

فئات العمر	ذكـــور	انـــاث
أقل من ٥	٦٦٧٧	٦٦١٦
٥ – ٩	٥٢٣٦	٥٢٢٨
١٠ – ١٤	٤٤٧٠	٣٥٧٧
١٥ – ١٩	٥١٧٧	٣٣٦١
٢٠ – ٢٤	١٠٣٠٨	٣٤٣٤
٢٥ – ٢٩	٩١٢٤	٣١٧٥
٣٠ – ٣٤	٦٣٨٠	٢٣٠٠
٣٥ – ٣٩	٤٣٠٠	١٩١١
٤٠ – ٤٤	٣٢٢٨	١٥١٨
٤٥ – ٤٩	٢٢٧٧	١١١٧
٥٠ – ٥٤	١٩٤١	١٢٤٣
٥٥ – ٥٩	٨٧٦	٦٢٨
٦٠ – ٦٤	٩٣٢	٧٩٤
٦٥ – ٦٩	٣٩٧	٣٠٩
٧٠ فاكثر	٨١٣	٨٨٣

ارسم هرم التركيب العمري وبين ملاحظاتك عليه .

٧ – جمعت البيانات الآتية عن ٥٠ مؤسسة تجـــارية . موقع المؤسسة ورأسمالها بالليرة والضرائب التي دفعتها في عام (بالليرة) .

بيروت	الجنوب	الشمال	بيروت
١ — ٢٠٠٠٠	٢ — ٣٠٠٠٠	٣ — ٢٥٠٠٠	٤ — ٥٠٠٠٠
٥٠٠	٧٠٠	٦٥٠	١٢٠٠
بيروت	الجنوب	الشمال	بيروت
٥ — ١٥٠٠٠	٦ — ١٠٠٠٠	٧ — ١٥٠٠٠	٨ — ٣٠٠٠٠
٤٥٠	٣٥٠	٤٠٠	٧٥٠
الشمال	الجنوب	بيروت	الشمال
٩ — ١٥٠٠٠	١٠ — ١٦٠٠٠	١١ — ٧٠٠٠٠	١٢ — ٣٥٠٠٠
٤٥٠	٥٠٠	١٥٠٠	٩٠٠
بيروت	الشمال	بيروت	الجنوب
١٣ — ٨٥٠٠٠	١٤ — ٦٥٠٠٠	١٥ — ٤٥٠٠٠	١٦ — ٤٠٠٠٠
١٧٠٠	١٢٠٠	٨٠٠	٧٥٠
بيروت	الشمال	الجنوب	بيروت
١٧ — ١٨٠٠٠	١٨ — ٣٥٠٠٠	١٩ — ٦٥٠٠٠	٢٠ — ٥٥٠٠٠
٤٥٠	٨٠٠	٩٠٠	٨٥٠
بيروت	الشمال	الجنوب	بيروت
٢١ — ٣٥٠٠٠	٢٢ — ١٢٠٠٠	٢٣ — ٧٥٠٠٠	٢٤ — ١٩٠٠٠
٥٠٠	٣٠٠	١١٠٠	٦٠٠
الشمال	الجنوب	الشمال	الجنوب
٢٥ — ٧٠٠٠٠	٢٦ — ٣٨٠٠٠	٢٧ — ٤٥٠٠٠	٢٨ — ١٨٠٠٠
٩٠٠	٨٠٠	٩٠٠	٦٠٠

بيروت	الشمال	بيروت	الشمال
٢٩— ٢٧٠٠٠	٣٠— ٥٥٠٠٠	٣١— ٩٠٠٠٠	٣٢— ٦٠٠٠٠
٧٥٠	٩٠٠	١٦٠٠	٩٥٠
الجنوب	بيروت	الشمال	بيروت
٣٣— ٤٨٠٠٠	٣٤— ١٧٠٠٠	٣٥— ٤٢٠٠٠	٣٦— ١٤٠٠٠
٧٥٠	٦٥٠	٩٥٠	٣٥٠
الشمال	بيروت	الجنوب	الشمال
٣٧— ١٨٠٠٠	٣٨— ٣٥٠٠٠	٣٩— ٧٣٠٠٠	٤٠— ١٦٠٠٠
٦٥٠	٩٥٠	١١٠٠	٥٥٠
بيروت	الشمال	الجنوب	بيروت
٤١— ٣٣٠٠٠	٤٢— ١٩٠٠٠	٤٣— ٢٣٠٠٠	٤٤— ١٧٠٠٠
٩٥٠	٥٥٠	٧٥٠	٦٠٠
الشمال	بيروت	الجنوب	بيروت
٤٥— ٣٨٠٠٠	٤٦— ٤٧.٠٠	٤٧— ٤٣٠٠٠	٤٨ - ١٣٠٠٠
٩٥٠	١٠٠٠	٩٥٠	٤٠٠
بيروت	الجنوب		
٤٩— ٢٧٠٠٠	— ١٨٠٠٠		
٨٠٠	٥٠٠		

المطلوب :

١ – تبويب المؤسسات التجارية ورؤوس اموالها والضرائب التي دفعتها تبعاً للموقع وفئات رأس المال بحيث يظهر الجدول المجاميع الخاصة بكل منطقة على حدة ومتوسطات رأس المال والضرائب .

الموقع – بيروت – الشمال – الجنوب

فئات رأس المال – ١٠٠٠٠ – ٢٠٠٠٠ – ٥٠٠٠٠ وأكثر

٢ – توضيح توزيع هذه المؤسسات تبعاً لموقعها برسم بياني :

٨ – كانت الارقام القياسية لنفقة المعيشة ولاسعار الجملة في احدى الدول في الفــترة : ١٩٥٢ – ١٩٦١ كالآتي (على أساس ١٩٣٩) :

السنوات	الرقم القياسي لنفقة المعيشة	الرقم القياسي لأسعار الجملة
١٩٥٢	٣١٧	٣٧٢
١٩٥٣	٢٩٦	٣٥٥
١٩٥٤	٢٨٤	٣٤٥
١٩٥٥	٢٨٣	٣٥١
١٩٥٦	٢٩٠	٣٨٩
١٩٥٧	٣٠٢	٤٢٢
١٩٥٨	٣٠٢	٤١٧
١٩٥٩	٣٠٣	٤١٧
١٩٦٠	٣٠٤	٤١٨
١٩٦١	٣٠٦	٤٢٥

وضح التغير في هذه الارقام القياسية في رسم بياني واحد .

٩ – البيانات الآتية عن عدد السكان وعدد الاسرة في احدى الدول في الفترة بين ١٩٥٢ – ١٩٦٢ :

السنة	عدد السكان بالألف	عدد الأسرة
١٩٥٢	٢١٤٧٣	٢٤٠٩٣
١٩٥٣	٢٢٠٠٣	٢٤٩٦٨
١٩٥٤	٢٢٥٥٧	٢٥٨٩٢
١٩٥٥	٢٣٠٦٣	٢٦٩٨٤
١٩٥٦	٢٣٦٤٣	٣٣٣٧٢
١٩٥٧	٢٤٢١٧	٣٠٤٠٧
١٩٥٨	٢٤٧٩١	٣٢١١٢
١٩٥٩	٢٥٣٦٥	٣٣١٥٣
١٩٦٠	٢٦٠٦٥	٣٧٠٩٤
١٩٦١	٢٦٢٧٤	٣٧٨٨٦
١٩٦٢	٢٧٠٦٤	٣٧٩٤٢

قارن التغير في عدد السكان وعدد الاسرة برسم بياني .

١٠ – البيانات الآتية تبين انتاج بعض المصنوعات في احدى الدول في عامي ١٩٥٢ ، ١٩٦٠ :

	١٩٥٢	١٩٦٠
عزل القطن –	٥٦ الف طن	١٠٥ الف طن
منسوجات قطنية –	٢٢٠ مليون متر مربع	٤٨٢ مليون متر مربع
زيت بذرة القطن –	١٠٦ الف طن	١٠٤ الف طن
نبيـذ –	١٦ الف طن	٣,٥ الف طن
سكر –	١٨٨ الف طن	٢٣٢ الف طن
اسمنت –	٩٤٧ الف طن	٢٠٤٧ الف طن

وضح التغير النسبي في الانتاج بين هذين العامين برسم بياني . احسب نسب التغير على أساس عام ١٩٥٢ .

١١ – الآتي ارقام قياسية لأسعار الجملة في دولة ما في خلال عام ١٩٥٨ على أساس ١٩٤٠ .

الشهر :	ينـاير	فبراير	مارس	ابريل	مـايو	يونيـو
الرقم القياسي :	١٠٩	٩٨	١٠٥	٩٦	٩٤	٩٣

الشهر :	يوليو	اغسطس	سبتمبر	اكتوبر	نوفمبر	ديسمبر
الرقم القياسي :	٩٣	٩٧	٩٦	١٩٤	١٠٣	١٠٨

وضح التغير في هذه الارقام .

١٢ – الآتي بيانات عن جملة قروض حكومة دولة مـا تبعاً لأنواعها المختلفة (الأرقام بملايين الليرات) .

السنة	قروض داخلية طويلة الأجـل	قروض داخلية قصيرة الأجل	قروض أجنبية
١٩٤٥	٦١٠٠	٨٠٠	١٠٠٠
١٩٤٦	٦٢٠٠	٧٠٠	١٠٠٠
١٩٤٧	٦٣٠٠	٨٠٠	١٠٠٠
١٩٤٨	٦٤٠٠	٩٠٠	١٠٠٠
١٩٤٩	٦٦٠٠	١٥٠٠	١٢٠٠
١٩٥٠	٧٧٠٠	٢٨٠٠	١٣٠٠
١٩٥١	٩٨٠٠	٣٣٠٠	١٥٠٠
١٩٥٢	١١٦٠٠	٤١٠٠	١٦٠٠
١٩٥٣	١٣٥٠٠	٥٠٠٠	١٧٠٠

وضح التغير الذي حدث في تكوين قروض هذه الحكومة برسم بياني .

١٣ – في دراسة عن المساكن ، استعملت الاستمارة الآتية :

١ – رقم المسكن المسلسل ____________

٢ – موقع المسكن المحافظة ________ القضاء ________

٣ – عدد الأسر التي تسكن في المسكن ____________

٤ – نوع بناء السكن طين ______ طوب ______
اسمنت ______ خيام ______

٥ – عدد الغرف ____________

٦ – هل السكن ملك ______ ايجار ______ هبة ______

٧ – مصدر المياه النقية مياه جارية ________ حنفية عامة ______
بئر خاص ________ بئر عام ________

٨ – اذا كان المسكن بالايجار كم الايجار السنوي ____________

٩ – عدد الافراد الذين يسكنون المسكن ________ منهم
________ ذكور و ______ اناث

منهم ________ فوق سن ١٣
و ________ بين ٥ سنوات و ١٣ سنة
و ________ أقل من ٥ سنوات

١٠ – طريقة تصريف القاذورات مجاري عامة ________
بئر ________ طريقة أخرى ________

١١ – نوع سقف البناء اسمنت ______ خشب ________
مواد أخرى ________

١٢ – مساحة المسكن ________ متر مربع

١٣ – هل هناك حديقة للمسكن نعم ______ لا ______

١٤ – متى اسس المسكن عام __________

المطلوب

١ – تصميم الدليل لترميز البيانات السابقة

٢ – تصميم الكشف الذي يمكن أن يستخدم لتوقيع الدليل عليه

٣ – افترض مسكن له اجابات معينة على هذه الاسئلة ورمّز هذه الاجابات في كشف الترميز الذي صممته .

١٤ – قامت دولة ما باجراء تعداد صناعي وكانت استمارة التعداد تحتوي على البيانات الآتية :

١ – اسم المؤسسة
٢ – موقع المؤسسة القضاء المحافظة
٣ – نشاط المؤسسة
٤ – الكيان القانوني للمؤسسة
٥ – عدد العمال الذكور
٦ – عدد العمال الاناث
٧ – مجموع عدد العمال
٨ – مجموع أجور الذكور في شهر
٩ – مجموع أجور الاناث في شهر
١٠ – المجموع العام للاجور
١١ – قيمة الانتاج

١٢ – قيمة المواد الخام الداخلة في الانتاج
١٣ – أنواع الآلات المستخدمة في الانتاج
١٤ – أنواع مولدات الطاقة المستخدمة
١٥ – قيمه الاستثمار الكلي في عام .

المطلوب

١ – اعداد الدليل الخاص بهذه البيانات

٢ – تصميم كشف توقع فيه أرقام الدليل ويجري التنقيب على أساسه .

١٥ – فيما يلي استمارة تشخيص خاصة بمرض البلهارسيا في أحد المستشفيات

اسم المريض .... رقم الادخال ....

عنوان المريض .... تاريخ الادخال ....

تاريخ الخروج ....

مدة البقاء في المستشفى ....

الجنس ذكر ——— انثى ———

العمر السنين ———الأشهر ———

المستوى الاقتصادي والاجتماعي

عالي ——— وسط ——— فقير ——— لم يبين ———

الديانة مسلم ——— مسيحي ——— ديانة اخرى ——— لم يبين ———

الوزن عند الدخول ——— كيلو ——— جرام

الوزن عند الخروج ——— كيلو ——— جرام

الطول ——— سم

نوع مكان السكن مدينة ——— قرية ——— بدو ———

مدينة قروية ——— لم يبين ———

المرض الأساسي أو الايذاء الذي ادى للادخال———

أحوال مرضية أخرى ساعدت (حادة أو مزمنة) ———

التشخيص

الاسنان	نعم ———	لا ———	لم يبين ———
كبد متمدد	نعم ———	لا ———	لم يبين ———
بنكرياس متمدد	نعم ———	لا ———	لم يبين ———
حمى	نعم ———	لا ———	لم يبين ———
اسهال حاد	نعم ———	لا ———	لم يبين ———
اسهال معاود	نعم ———	لا ———	لم يبين ———
اسهال مع صديد	نعم ———	لا ———	لم يبين ———
اسهال مع دم	نعم ———	لا ———	لم يبين ———
اسهال مع نقط دهنية	نعم ———	لا ———	لم يبين ———
اسهال غير محدد	نعم ———	لا ———	لم يبين ———
جلد مقشر	نعم ———	لا ———	لم يبين ———

سبب الخروج من المستشفى :

العلاج انتهى ——— خروج بدون اذن ——— حول الى مستشفى

آخر ——— الموت ——— سبب آخر———

لم يبين ———

الحالة عند الخروج .

تحسن ———— لم يتغير ———— أسوأ ———— وفاة————

لم يبين ————

المطلوب :

١ – تصميم الدليل الخاص بترميز البيانات السابقة .

٢ – تصميم بطاقة التثقيب الخاصة بهذا الدليل .

٣ – تصميم كشف للترميز مع مثال .

١٦ ـ الأرقام الآتية تبين أطوال وأوزان مجموعة من الطلبة :

الطول سم	الوزن كيلو	الطول سم	الوزن كيلو	الطول سم	الوزن كيلو
١٧٠	٧٥	١٥٥	٦٠	١٥٤	٥٧
١٥٥	٦٠	١٤١	٥٢	١٦٨	٧٥
١٥٦	٦٣	١٤٧	٥٥	١٦٩	٨٢
١٤٣	٥٥	١٦٦	٧٠	١٥٥	٦٥
١٤٩	٥٩	١٦٩	٧٩	١٦٣	٦٣
١٦٨	٧٨	١٥٢	٥٤	١٥٤	٥٩
١٤٦	٦٢	١٧١	٨١	١٦٩	٧٦
١٥٣	٥٧	١٦٣	٥٩	١٧٣	٧٨
١٦٤	٦٢	١٧٣	٧٩	١٤٤	٥٥
١٦٩	٧٩	١٤٨	٥٧	١٦٥	٧٧
١٤٨	٥١	١٥٢	٥٥	١٦٩	٧٤
١٦٦	٦٦	١٧٤	٨٠	١٤٨	٥٧
١٥٤	٥٣	١٦٦	٧٥	١٦٣	٦٩
١٦٤	٦٩	١٥٣	٦٤	١٧١	٧٣
١٤٤	٥١	١٦٧	٧٤	١٦٣	٦٨
١٤٠	٥٩	١٦٨	٧٦	١٦٦	٧٥
١٤٢	٥٨	١٤٣	٥٠	١٥٧	٦٤
١٥٦	٦١	١٦١	٦٨	١٥٩	٦٦
١٦٣	٦٧	١٧٤	٧٨	١٥٦	٦٧
١٦٥	٦٩	١٧٣	٧٦	١٤٣	٥٩

٢ ـ المطلوب ترتيب هذه البيانات في جدول تكراري مزدوج .

١٧ – الآتي توزيع أوزان مجموعة من الطلبة :

الوزن بالكيلو	عدد الطلبة
٤٥	٢
٥٠	٣
٥٥	٣
٦٠	٦
٦٥	١٢
٧٠	٢٣
٧٥	٣١
٨٠	١٨
٨٥	١٨
٩٠	١٠
٩٥	٦

المطلوب :

١ – توضيح هذا التوزيع بهستو جرام .
٢ – توضيح هذا التوزيع بمضلع تكراري .
٣ – توضيح هذا التوزيع بمنحنى تكراري .
٤ – توضيح هذا التوزيع بمنحنى متجمع صاعد وهابط .
٥ – تحديد عدد الطلبة الذين يقل وزنهم عن ٨٦ كيلو .
٦ – تحديد الوزن الذي يقل عنه نصف أفراد المجموعة .
٧ – تحديد الوزن الذي يقل عنه ربع أفراد المجموعة .
٨ – تحديد عدد الطلبة الذين تتراوح أوزانهم بين ٦٢ و ٦٨ كيلو .
٩ – تحديد عدد الطلبة الذين تكون أوزانهم ٧٢ كيلو وأكثر .

مقاييس التركز

الوسط الحسابي والهندسي والتوافقي والمنوال والوسيط

الفصل الخامس

المتوسطـــات

بعد تلخيص البيانات في توزيع تكراري (بالنسبة للظـــواهر القيمية المتغيرة) ننتقل إلى التحليل الاحصائي لهذا التوزيع الذي يتكون أساساً من ثلاث أبحاث هي :

١ — البحث عن القيمة التي تتركز حولها القيم التي تتخذها الظاهرة في تغيرها من وحدة الى اخرى (المتوسط) .

٢ — البحث عن مدى تشتت القيم حول هذه القيمة المتوسطة (التشتت).

٣ — البحث عن تماثل التوزيع حول القيمة المتوسطة (الالتواء) .

وفي هذا الفصل سنناقش مقاييس التركز التي جرى العرف الاحصائي على تسميتها بالمتوسطات وهي تشمل الوسط الحسابي والهندسي والتوافقي والمنوال والوسيط وغير ذلك من المتوسطات التي لن نتعرض لهـا في هذا الكتاب .

والفكرة الاساسية التي يقوم عليها موضوع المتوسطات هي أن تمثيل التوزيع التكراري بقيمة واحدة يبرره ميل المجموعات الكبيرة من الوحدات نحو أن تتركز قيمها حول قيمة معينة تنحرف عنها القيم الاخرى بشيء من

الانتظام ، هذه القيمة هي التي نسميها بالمتوسط وان كانت تتخذ أسماء مختلفة تبعاً للاساس الذي ينبني عليه قياسها . وبذلك يكون أي متوسط هو القيمة التي نعتبرهـــا ممثلة لمجموعة القيم التي حسب لها هذا المتوسط . وإذا كانت الانواع المختلفة من المتوسطات تعتبر ممثلة للقيم التي حسبت لهـــا مع اختلاف قيمة كل نوع منها، فليس في ذلك أي تناقض حيث أن هناك أسباب معينة خاصة بكل متوسط والتي على اساسها يعتبر ممثلاً للقيم التي حسب لها وسوف ندرس هذه الاسباب عندما نأتي إلى مناقشة كل متوسط على حدة .

الوسط الحسابي :

نفترض ان لدينا ظاهرة معينة س تتغير قيمتها من وحدة الى اخرى (مثلاً أجور مجموعة من العمال) وبذلك يصبح لدينا القيم $س_1$ ، $س_2$ ، $س_3$ ، $س_4$، ..... $س_ن$ ، فإن كان عدد هذه القيم (عدد الوحدات موضوع الدرس) ن يكون الوسط الحسابي لهذه القيم :

$$\bar{س} = \frac{س_1 + س_2 + س_3 + س_4 + \cdots + س_ن}{ن}$$

اى ان $\bar{س} = \frac{\text{مج}\ س}{ن}$ (الرمز مج يدل على مجموع القيم)

وبذلك يكون مج س = $ن \times \bar{س}$

ما هي خائص هذه القيمة $\bar{س}$ حتى نعتبرها ممثلة للقيم التي حسبت لها ؟ ان عملية حساب هذه القيمة تجعلها توازن بين القيم الصغيرة والقيم الكبيرة من حيث انحرافها عنها ؛ بمعنى ان مجموع انحرافات القيم الكبيرة عن $\bar{س}$ = مجموع انحرافات القيم الصغيرة عن $\bar{س}$ ، أي ان المجموع الجبري لانحرافات جميع القيم عـن الوسط الحسابي = صفر ولاثبات ذلك نحسب انحرافات القيم عن $\bar{س}$ ونجمعها كالآتي :

$$(س_١ - \bar{س}) + (س_٢ - \bar{س}) + (س_٣ - \bar{س})$$
$$(س_٤ - \bar{س}) \cdots (س_ن - \bar{س})$$

بفك الأقواس $= مجـ\, س - ن\, \bar{س}$

$= ن\, \bar{س} - ن\, \bar{س} =$ صفر .

والخاصية الثانية للقيمة $\bar{س}$ هي ان مجموع مربع انحرافات القيم عنها أصغر من مجموع مربع انحرافات القيم عن أي قيمة أخرى .

على أساس هاتين الخاصيتين للوسط الحسابي تعتبر قيمته ممثلة لجميع القيم التي حسب لها .

طريقة أخرى لحسلب الوسط الحسابي :

في المثال السابق اذا أخذنا قيمة معينة وافترضنا انها الوسط الحسابي ورمزنا لهذه القيمة بالرمز و وحسبنا انحرافات القيم عن و وظهر ان مجموع الانحرافات = صفر تكون هذه القيمة هي نفسها الوسط الحسابي ، أما إذا ظهر ان مجموع الانحرافات لا يساوي صفراً فيجب تصحيح هذه القيمة تصحيحاً معيناً حتى نجعلها تساوي $\bar{س}$ أي حتى نجعل مجموع الانحرافات عنها = صفر . نحسب الآن انحرافات القيم عن و ونجمع هذه الانحرافات .

$$(س_١ - و) + (س_٢ - و) + (س_٣ - و) + (س_٤ - و) +$$
$$\cdots (س_ن - و)$$

وبفك الأقواس $= مجـ\, س - ن\, و =$ مجموع الانحرافات عن الوسط الفرضي .

$= ن\, \bar{س} - ن\, و =$ مجموع الانحرافات عن الوسط الفرضي .

وبالقسمة على ن يكون $= \bar{س} - و = \dfrac{\text{مجموع الانحرافات عن الوسط الفرضي}}{ن}$

فاذا رمزنا لمجموع الانحرافات عن الوسط الفرضي بالرمز مجـ ح

$$\therefore \bar{س} = و + \frac{\text{مجـ ح}}{ن}$$

الوسط الحسابي المرجح :

نلاحظ انه في المثال السابق كانت القيم متعادلة في الاهمية ، لكن نفترض أن القيمة $س_1$ تكررت $ك_1$ مرة والقيمة $س_2$ تكررت $ك_2$ مرة والقيمة $س_3$ تكررت $ك_3$ وهكذا ، وان $ك_1$ و $ك_2$ و $ك_3$ غير متساوية هل يحسب الوسط الحسابي بنفس الطريقة السابقة . لا شك ان استعمال نفس الطريقة السابقة يجعلنا نحصل على قيمة مضللة بعض الشيء اذ انه يجب أن يؤخذ عدد مرات تكرار كل قيمة في حسابنا حتى نحصل على وسط حسابي غير مضلل . ان عدد مرات تكرار القيم يدل على اهميتها ولذلك نسميها بالاوزان والوسط الحسابي لهـذا النوع من القيم نسميه الوسط الحسابي المرجح حيث اننا نرجح كل قيمة بوزنها الذي يدل على أهميتها ويكون ذلك كالآتي :

$$\bar{س} = \frac{س_1 \times ك_1 + س_2 \times ك_2 + س_3 \times ك_3 + س_4 \times ك_4}{ك_1 + ك_3 + ك_2 + ك_4}$$

والوسط المحسوب بهذه الطريقة يتوفر فيه كذلك الخاصيتان التي سبق الاشارة اليها ، حيث اننا في الواقع لم نغير شيئاً في طريقة الحساب اذ بدل أن نكرر $س_1$ جمعاً $ك_1$ مرة ضربناها في $ك_1$ وكذلك بدل أن نكرر $س_2$

جمعاً $ك_٢$ مرة ضربناها في $ك_٢$ ، أما المقام $ك_١ + ك_٢ + ك_٣ + ك_٤$ فهذا ايضاً يدل على العدد الكلي للقيم الموجودة في البسط .

والآن نحسب انحرافات القيم عن وسطها الحسابي ونجمعها كالآتي :

$(س_١ - \bar{س}) ك_١ + (س_٢ - \bar{س}) ك_٢ + (س_٣ - \bar{س}) ك_٣$ وهكذا.

$= مجـ س - \bar{س} (ك_١ + ك_٢ + ك_٣)$

$= مجـ س - ن \bar{س}$

$= مجـ س - مجـ س =$ صفر

حيث ان $س_١ ك_١ =$ مجموع القيمة الاولى و$س_٢ ك_٢ =$ مجموع القيمة الثانية و $س_٣ ك_٣ =$ مجموع القيمة الثالثة ، وبذلك يكون مجموعها جميعاً مساوياً للمجموع الكلي لقيم الظاهرة س .

مثال ١ :

نفترض ان لدينا خمس عمال أجورهم كالآتي ٨ ، ٣ ، ٥ ، ٩ ، ١٠

يكون $\bar{س} = \frac{٣٥}{٥} = ٧$

نحسب الآن انحرافات القيم عن ٧ ونجمعها : —

+ ١ — ٤ — ٢ + ٢ + ٣ = صفر

نربع هذه الانحرافات نحصل على ١ + ١٦ + ٤ + ٤ + ٩ = ٣٤

إذا أخذنا قيمة أخرى فرضية ٦ ونحسب الانحرافات عنها

+ ٢ — ٣ — ١ + ٣ + ٤ = ٥ (لا يساوي صفر)

نربع هذه الانحرافات نحصل على ٤ + ٩ + ١ + ٩ + ١٦ = ٣٩

وهو أكبر من مجموع مربع الانحرافات عن الوسط الحسابي .

يمكننا أن نحصل على الوسط الحسابي عن طريق هذه القيمة الفرضية ٦ كالآتي : —

$$\bar{\text{س}} = \text{و} + \frac{\text{مجـ ح}}{\text{ن}}$$

$$= ٦ + \frac{٥}{٥} = ٧$$

مثال ٢ :

نفترض أن ١٠٠٠ عامل يحصل كل منهم على أجر ٥ ليرات .

وان ١٠ عمال يحصل كل منهم على أجر ١٠٦ ليرات .

اذا حسبنا المتوسط كالآتي $\frac{١٠٦ + ٥}{٢}$ = ٥٥,٥ نلاحظ انه قيمة مضللة حيث أن معظم العمال يتقاضون في الواقع أجراً لا يزيد عن ٥ ليرات .

ولكن بحساب الوسط المرجح $\frac{١٠٠٠ \times ٥ + ١٠ \times ١٠٦}{١٠١٠} = \frac{١٠٦٠٠}{١٠١٠} = ٦$

وهو أجر قريب جداً من أجر غالبية العمال :

نحسب انحرافات القيم من المتوسط المرجح ٦ ونجمع هذه الانحرافات :

(−١) ١٠٠٠ + (١٠٠) ١٠ = صفر .

اذا أخذنا قيمة قيمة أخرى فرضية ٩ ونحسب الانحراف عنها :

$$(-٤) \times ١٠٠٠ + (٩٧) ١٠ = - ٣٠٣٠$$

الوسط الحسابي $\bar{س} = و + \frac{مج\ ح}{ن}$

$$= ٩ + \frac{-٣٠٣٠}{١٠١٠}$$

$$= ٩ - ٣ = ٦$$

الوسط الحسابي لتوزيع تكراري :

لحساب الوسط الحسابي لتوزيع تكراري يمكن أن تتبع الخطوات الآتية :

١ – نحسب مركز كل فئة حيث أن القيمة المتوسطة للفئة .

٢ – نضرب مركز كل فئة في تكرارها ونجمع النتائج .

٣ – نستعمل القانون $\bar{س} = \frac{مج\ م\ ك}{مج\ ك}$ (نلاحظ ان هذه الطريقة لا لا تختلف عن طريقة الوسط الحسابي المرجح التي سبق الاشارة اليهـا وليس هناك من فرق الا اننا عبرنا عن القيم بالرمز م بدلا من الرمز س) .

مثال ٣ : (توزيع تكراري غير متصل)

الفئات	التكرارات(ك)	المراكز (م)	المراكز التكرارات (م ك)
٦٠–٦٢	٥	٦١	٣٠٥
٦٣–٦٥	١٨	٦٤	١١٥٢
٦٦–٦٨	٤٢	٦٧	٢٨١٤
٦٩–٧١	٢٧	٧٠	١٧٨٠
٧٢–٧٤	٨	٧٣	٥٨٤
	١٠٠		٦٧٤٥

$$\bar{س} = \frac{٦٧٤٥}{١٠٠} = ٦٧٫٤٥$$

(نلاحظ اننا حسبنا مركز الفئة الأولى ٦١ وأضفنا باستمرار ٣ الذي هو مدى الفئة كي نحصل على مراكز الفئات التالية) .

مثال ٤ : (توزيع تكراري متصل)

ف	ك	م	م × ك
٦٠ —	٥	٦١٫٥	٣٠٧٫٥
٦٣ —	١٨	٦٤٫٥	١٩٦١
٦٦ —	٤٢	٦٧٫٥	٢٨٣٥
٦٩ —	٢٧	٧٠٫٥	١٩٠٣٫٥
٧٢ — ٧٥	٨	٧٣٫٥	٥٨٨
	١٠٠		٦٧٩٥

$$\bar{س} = \frac{٦٧٩٥}{١٠٠} = ٦٧٫٩٥$$

مثال ٥ : (توزيع تكراري غير منتظم)

ف	ك	م	م × ك
٥ —	٥	٧٫٥	٣٧٫٥
١٠ —	١٨	١٥	٢٧٠
٢٠ —	٤٢	٢٥	١٠٥٠
٣٠ —	٢٧	٤٠	١٠٨٠
٥٠ — ٧٠	٨	٦٠	٤٨٠
	١٠٠		٢٩١٧٫٥

$$\bar{س} = \frac{٢٩١٧,٥}{١٠٠} = ٢٩,١٧٥$$

ان استخـــدام الطريقة السابقة $\bar{س} = \frac{مج\ م\ ك}{مج\ ك}$ يستدعي عمليـــات حسابية طويلة يضيع فيها الوقت ولذلك يكون من الأفضل أن نلجأ إلى الطريقة الثانية التي يحسب فيهـــا الوسط بناء على قيمة نفترضها وبذلك نستعمل القانون :

$$\bar{س} = و + \frac{مج\ ح\ ك}{مج\ ك}$$

ونلاحظ ان هذا القانون لا يختلف عن القانون الذي سبق الاشارة اليه فالقيمة مج ح ك تدل على مجموع الانحرافات عن الوسط الفرضي وادخلنا ك في الحساب لأن انحرافات القيم عن الوسط الفرضي تتكرر تبعاً لتكرار القيمة نفسها . كذلك يكون مج ك هو نفسه عبارة عن عدد القيم .

مثال ٦ :

ف	ك	م	ح	ح ك
٦٠ — ٦٢	٥	٦١	— ٦	— ٣٠
٦٣ — ٦٥	١٨	٦٤	— ٣	— ٥٤
٦٦ — ٦٨	٤٢	٦٧	صفر	
٦٩ — ٧١	٢٧	٧٠	+ ٣	+ ٨١
٧٢ — ٧٤	٨	٧٣	+ ٦	+ ٤٨
	١٠٠			٤٥

$\bar{س} = ٦٧ + \frac{٤٥}{١٠٠}$

$$\therefore\ \bar{\text{س}} = ٦٧ + \frac{٤٥}{١٠٠} = ٦٧{,}٤٥$$

ونلاحظ ان الانحرافات ٦ ، — ٣ ، + ٣ ، + ٦ تختصر جميعاً على طول الفئة ٣ ولذلك تسهيلا للعمل نختصر هذه الانحرافات وفي هذه الحالة نستخدم القـــانون :

$$\bar{\text{س}} = \text{و} + \frac{\text{مجـ ح ك}}{\text{مجـ ك}} \times \text{طول الفئة}$$

نلاحظ اننا نضرب مجـ ح ك × طول الفئة حيث اننا عند ايجاد مجموع الانحرافات عن الوسط الفرضي اختصرناها جميعاً على طول الفئة ولذلك نعود فنضربها في طول الفئة حتى نرجعها الى أصلها ، أما إذا كانت الفئات غـــير منتظمة نحسب الانحرافات الواقعية عن الوسط الفرضي ولا يكون هناك داعي للضرب في طول الفئة عندتطبيق القانون .

مثال ٧ :

ف	ك	م	ح	ح ك
٦٠ — ٦٢	٥	٦١	— ٢	— ١٠
٦٣ — ٦٥	١٨	٦٤	— ١	— ١٨
٦٦ — ٦٨	٤٢	٦٧	صفر	صفر
٦٩ — ٧١	٢٧	٧٠	+ ١	+ ٢٧
٧٢ — ٧٤	٨	٧٣	+ ٢	+ ١٨
	١٠٠			١٥

$$\therefore \bar{س} = ٦٧ + \frac{١٥}{١٠٠} \times ٣ = ٦٧ر٤٥$$

الوسيط :

الوسيط لأي مجموعة من القيم هو القيمة التي تقع في الوسط أي التي يقل عنها عدد من القيم يساوي عدد القيم التي تزيد عنها ، وهذه هي الخاصية التي على اساسها يكون الوسيط هو القيمة الممثلة لمجموعة القيم التي حسب لهـا . فالوسيط هو القيمة التي توازن بين القيم الصغيرة والقيم الكبيرة من حيث عددها .

ولحساب الوسيط يجب ترتيب القيم ترتيباً تصاعدياً حتى اذا حسبناه في المنتصف تكون القيم السابقة له أقل منه والقيم اللاحقة له أكـبر منه . بعد ترتيب القيم نجد ترتيب الوسط ثم نبحث عن القيمة المقابلة لهذا الترتيب، وإذا كان ترتيب الوسيط يقع بين قيمتين نجد الوسط الحسابي لهما .

مثال ٨ :

نفترض ان لدينا القيم ٧٨ ، ٧٢ ، ٦٨ ، ٨٧ ، ٨٤ ، ٩١ ، ونريد أن نحسب لها الوسيط . نرتب القيم ترتيباً تصاعدياً كالآتي :

٦٨ ، ٧٢ ، ٧٨ ، ٨٤ ، ٨٧ ، ٩١

نحسب بعد ذلك ترتيب الوسيط وهو يساوي $\frac{\text{عدد القيم} + ١}{٢} =$

$\frac{٦ + ١}{٢} =$ ٣ر٥ اي ان الوسيط هو القيمة الثالثة والنصف وهذا الترتيب يقع بين ٧٨ ، ٨٤ .

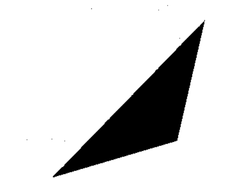
$$\therefore \text{الوسيط} = \frac{٧٨ + ٨٤}{٢} = \frac{١٦٢}{٢} = ٨١$$

وعندما تكون القيم مبوبة في توزيـع تكراري يجب اولا وضع التوزيـع في شكل متجمع صاعداً أو هابط (يكون ذلك ترتيباً للقيم) ثم نحدد ترتيب الوسيط وهو يساوي مجموع التكرارات مقسوماً على ٢ ثم نحاول ان نجد القيمة المقابلة لهذا الترتيب ولتوضيح ذلك نفترض المثال الآتي :

مثال ٩ :

ف	ك	الحدود العليا للفئات	التكرار المتجمع الصاعد
١١٨ –	٣	أقل من ١٢٧	٣
١٢٧ –	٥	« « ١٣٦	٨
١٣٦ –	٩	« « ١٤٥	١٧
١٤٥ –	١٢	« « ١٥٤	٢٩
١٥٤ –	٥	« « ١٦٣	٣٤
١٦٣ –	٤	« « ١٧٢	٣٨
١٧٢ – ١٨١	٢	« « ١٨١	٤٠
	٤٠		

بـــين التكرار الصاعد ١٧ والتكرار
يط وهي للقيمة المقابلة لهذا الترتيب تقع
تغير في التوزيـع هو تغير منتظم يكون
مستقيماً ويكون ميله بين هاتين القيمتين



$\frac{٢٩ - ١٧}{١٥٤ - ١٤٥}$ وحيث انه خط مستقيم يكون هذا الميل مساوياً للميل عند النقطة التي تمثل ترتيب الوسيط ٢٠

اي : $\frac{٢٠ - ١٧}{س - ١٤٥}$ (س تدل على قيمة الوسيط)

∴ $\frac{١٢}{٩} = \frac{٣}{س - ١٤٥}$ ∴ ١٢ س = ١٧٤٠ = ٢٧

∴ ١٢ س = ١٧٦٧

س (الوسيط) = ١٤٧,٢٥ .

مثال ١٠ :

ف	ك	الحدود العليا للفئات	التكرار المتجمع الصاعد
١٠ −	٣	أقل من ١٥	٣
١٥ −	٧	» » ٢٠	١٠
٢٠ −	١٦	» » ٢٥	٢٦
٢٥ −	١٢	» » ٣٠	٣٨
٣٠ −	٩	» » ٣٥	٤٧
٣٥ −	٥	» » ٤٠	٥٢
٤٠ − ٤٥	٢	» » ٤٥	٥٤
	٥٤		

ترتيب الوسيط = $\frac{٥٤}{٢}$ = ٢٧ ويقع بين التكرار المتجمع الصاعد ٢٦ والتكرار الصاعد ٣٨ ، وتكون قيمة الوسيط هي القيمة المقابلة لهذا الترتيب وهي تقع بين ٢٥ ، ٣٠ .

$$\therefore \frac{٣٨ - ٢٦}{٣٠ - ٢٥} = \frac{٢٧ - ٢٦}{س - ٢٥}$$

$$\therefore \frac{١٢}{٥} = \frac{١}{س - ٢٥}$$

$$\therefore ١٢ \text{ س} - ٣٠٠ = ٥$$

$$\therefore ١٢ \text{ س} = ٣٠٥$$

$$\therefore \text{س (الوسيط)} = \frac{٣٠٥}{١٢} = ٢٥ \text{ ر } ٤١$$

ونلاحظ ان قيمة الوسيط تقع في الفئة ٢٥ – ٣٠ حيث ان عدد التكرارات السابقة لهذه الفئة ٣ + ٧ + ١٦ = ٢٦ أما ترتيب الوسيط فهو ٢٧ وبذلك تكون قيمته واقعة في هذه الفئة وحتى تحصل على هذه القيمة نحسب ما يقابل الزيادة في ترتيب الوسيط عن مجموع التكرارات السابقة (٢٧ – ٢٦) من طول الفئة ٥ .

وذلك يساوي $\frac{١}{١٢}$ × ٥ = ٤٢ر٠ ∴ الوسيط = ٢٥ + ٤٢ر٠ = ٢٥ر٤٢ .

الربيع الأول هو القيمة التي يسبقها ربع عدد القيم ويزيد عنها ثلاث أرباع القيم ، الربيع الثالث فهو القيمة التي يسبقها

ويمكن استخدام الطرق السابقة في ايجـــاد جميع القيم الموضعية الاخرى مثل الربيع الأدنى أو الربيع الاول والربيع الاعلى أو الثالث والعشير الاول والعشير الثاني وهكذا . والربيع الاول هو القيمة التي يسبقها ربع عدد القيم ويزيد عنها ثلاث أرباع عدد القيم . أما الربيع الثالث فهو القيمة التي يسبقها ثلاث أرباع عدد القيم ويزيد عنها ربع عدد القيم . والعشير الاول هو القيمة التي يقل عنها عشر عدد القيم ويزيد عنها تسع أعشار عدد القيم .

العشير الأول هو القيمة التي يقل عنها عشر عدد القيم ويزيد عنها تسع أعشار القيم

مثال ١١ :

ف	ك	الحدود العليا للفئات	التكرار المتجمع الصاعد
٣٠ –	١	اقل من ٤٠	١
٤٠ –	٣	اقل من ٥٠	٤
٥٠ –	١١	اقل من ٦٠	١٥
٦٠ –	٢١	اقل من ٧٠	٣٦
٧٠ –	٤٣	اقل من ٨٠	٧٩
٨٠ –	٣٢	أقل من ٩٠	١١١
٩٠–١٠٠	٩	اقل من ١٠٠	١٢٠

ترتيب الوسيط = $\frac{١٢٠}{٢}$ = ٦٠ واقع بين ٣٦ ، ٧٩

$$\therefore \frac{٧٩ - ٣٦}{٨٠ - ٧٠} = \frac{٦٠ - ٣٦}{س - ٧٠}$$

$$\therefore \frac{٤٣}{١٠} = \frac{٢٤}{س - ٧٠}$$

∴ ٤٣ س − ١٠١٠ = ٢٤٠

∴ ٤٣ س = ٣٢٥٠

∴ س = ٧٥,٥

ترتيب الربيع الأعلى = $\frac{\text{١٢٠}}{\text{٤}}$ = ٣٠ واقع بين ١٥ ، ٣٦

∴ $\frac{\text{٣٦ − ١٥}}{\text{٧٠ − ٦٠}}$ = $\frac{\text{٣٠ − ١٥}}{\text{س − ٦٠}}$

∴ = $\frac{\text{٢١}}{\text{١٠}}$ = $\frac{\text{١٥}}{\text{س − ٦٠}}$

∴ ٢١ س − ١٢٦٠ = ١٤٠

∴ ٢١ س = ١٤١٠

∴ س = ٦٧,١

ترتيب الربيع الاعلى : $\frac{\text{١٢٠}}{\text{٤}}$ × ٣ = ٩٠ واقع بين ٧٩ ، ١١١

∴ $\frac{\text{١١١ − ٧٩}}{\text{٩٠ − ٨٠}}$ = $\frac{\text{٩٠ − ٧٩}}{\text{س − ٨٠}}$

∴ $\frac{\text{٢٣}}{\text{١٠}}$ = $\frac{\text{١١}}{\text{س − ٨٠}}$ ∴ ٣٢ س − ٢٥٦٠ = ١١٠

∴ ٣٢ س = ٢٦٧٠

س = ٨٣,٤

ويمكن تحديد قيمة الوسيط وغيره من القيم الموضعية بواسطة الرسم البياني ، وذلك برسم المنحنى المتجمع الصاعد أو الهابط من واقع التوزيع ، ثم نحدد موقع ترتيب الوسيط أو أي قيمة موضعية أخرى على المحور الرأسي ، ثم نصل النقطة الدالة على هذا الموقع مع المنحنى المتجمع ومن نقطة التقاطع نسقط عمود على المحور الافقي يقابله في نقطة تكون القيمة عندها هي الوسيط أو القيمة الموضعية التي نبحث عنها . واذا رسمنا المنحنى الصاعد والهابط في رسم واحد فانهما يتقاطعان في نقطة تحدد ترتيب الوسيط ومنهما ننزل عمود على المحور الأفقي يقابله في قيمة تكون هي الوسيط .

المنوال :

المنوال لمجموعة من القيم هو القيمة الاكثر انتشاراً أو الأكثر تكراراً بين القيم ، وبمعنى آخر هي القيمة الشائعة ، وهذا هو الأساس الذي بناء عليه يعتبر المنوال وسطاً ممثلاً للقيم التي حسب لأجلها . إلا أن هذه القيمة قد لا توجد وحتى عند وجودها قد لا تكون قيمة ذات دلالة فعلية .

مثال ١٢ :

لمجموعة القيم ٢ ، ٢ ، ٥ ، ٧ ، ٩ ، ١ ، ٩ ، ١ ، ١١ ، ١٢ ، ١٨
المنوال هو ٩ .

لمجموعة القيم : ٣ ، ٥ ، ٨ ، ١ ، ٢ ، ١ ، ١٢ ، ١٥ ، ١٦
لا يوجد منوال .

لمجموعة القيم : ٢ ، ٣ ، ٤ ، ٤ ، ٤ ، ٥ ، ٥ ، ٧ ، ٧ ، ٧ ، ٩ هناك منوالان ٤ ، ٧ .

وبالنسبة للتوزيع التكراري لا يكون من السهل تحديد قيمة المنوال . وسنناقش هنا طريقتين لايجاد المنوال ولكن يجب أن نتذكر دائماً ان كلتا الطريقتين تقريبيتين . والخطوة الاولى في تحديد المنوال هي تحديد الفئة التي يكون تكرارها أكبر تكرار في التوزيع الموجود لدينا ، وواضح اذا غيرنا هذا التوزيع باستخدام فئات ذات مدى مختلف فسوف نحصل على فئة منوالية اخرى أي فئة يكون تكرارها هو الاكبر وبذلك تختلف قيمة المنوال بعض الشيء تبعاً لاختلاف مدى الفئة المستخدم في التوزيع ، ويعني ذلك ان قيمة المنوال ليست قيمة مستقرة وذلك هو السبب في عدم دقة قيمة المنوال كما سبق ان ذكرنا . وتنشأ هذه الصعوبة غالباً (صعوبة تحديد قيمة المنوال) بسبب تبويب البيانات في فئات واسعة المدى ، الأمر الذي يضيع معالم التوزيع ، أما اذا بوبنا البيانات في فئات ضيقة المدى فان قيمة المنوال تستقر الى حد ما عند رقم معين .

على انه يمكن بتمهيد المنحنى التكراري الذي يمثل التوزيع أن نحصل على قيمة دقيقة نوعاً ما للمنوال الذي يكون هو القيمة المقابلة لقمة المنحنى ، حيث ان هذه القيمة سوف تمثل أكبر تكرار ، وذلك لأن تمهيد المنحنى يكون بمثابة وضع التوزيع في فئات ضيقة المدى . إلا أن هذه الطريقة بها عيب يرجع الى صعوبة تمهيد المنحنى التكراري تمهيداً دقيقاً خاصة وان ذلك يتوقف الى حد كبير على خبرة الباحث بهذا الموضوع .

والطريقة الحسابية الاولى لايجاد المنوال تسمى طريقة الرافعة وتنحصر في تحديد الفئة المنوالية للتوزيع وهي الفئة ذات التكرار الاكبر ، كما قدمنا، فيكون المنوال هو الحد الادنى لهذه الفئة مضافاً اليه جزء من طول الفئة المنوالية يتناسب مع التكرارات السابقة واللاحقة لهذهالفئة والتي تعتبر كقوى مؤثرة على موضع المنوال فيها .

مثـال ١٣ :

ف	ك	
صفر –	٢٦٢	
١٠ –	٣٣٣	الفئة المنوالية
٢٠ –	٢١٠	
٣٠ – ٤٠	١٤٢	

∴ المنوال = ١٠ + (٢٠ − ١٠) × $\frac{٢١٠}{٢١٠ + ٢٦٢}$

= ١٠ + ٤٫٤٤٩ = ١٤٫٤٤٩

ولتوضيح الاجابة نفترض ان الفئة المنوالية هي كرافعة يؤثر فيها القوتان ٢٦٢ ، ٢١٠ ويكون توازنها كالآتي :

١٠ س | ١٠ − س ٢٠

٢٦٢ ↓ ↓ ٢١٠

∴ ٢٦٢ س = ٢١٠ (١٠ − س)

∴ ٢٦٢ س = ٢١٠٠ − ٢١٠ س

∴ ٤٧٢ س = ٢١٠٠

س = $\frac{٢١٠٠}{٤٧٢}$ = ٤٫٤٤٩

∴ المنوال = ١٠ + ٤٫٤٤٩ = ١٤٫٤٤٩

ويمكن وضع هذه الطريقة في قانون كالآتي :

المنوال = الحد الادنى للفئة المنوالية + $\frac{\text{التكرار اللاحق لهذه الفئة}}{\text{التكرار السابق+التكرار اللاحق}}$ × طول الفئة

والطريقة الثانية لحساب المنوال تسمى طريقة الفروق أو طريقة بيرسون تبعاً لاسم صاحبها وتقوم على أساس تحديد الفئة المنوالية كالطريقة السابقة تماماً ولكنها تختلف عنها في ان الذي يؤثر على موضع المنوال في هذه الفئة ليس هو التكرار السابق والتكرار اللاحق وإنما الفرق بين تكرار الفئة المنوالية وتكرار الفئتين السابقة واللاحقة لها . ويكون تحديد موضع المنوال في الفئة المنوالية على أساس تقسيمها بنسبة الفرق السابق الى الفرق اللاحق .

مثال ١٤ :

ف	ك	
صفر –	٢٦٢	
١٠ –	٣٣٣	الفئة المنوالية
٢٠ –	٢١٠	
٣٠ – ٤٠	١٤٢	

والمنوال بهذه الطريقة = ١٠ + $\frac{\text{٣٣٣ − ٢٦٢}}{\text{(٣٣٣ − ٢٦٢) + (٣٣٣ − ٢١٠)}}$ ×١٠

= ١٠ + $\frac{\text{٧١}}{\text{٧١ + ١٢٣}}$ × ١٠

= ١٠ + ٣٫٦٦ = ١٣٫٦٦

ولتوضيح الاجابة نقسم الفئة المنوالية تقسيماً جبرياً بنسبة س : ١٠ — س وحسابياً بنسبة الفرق السابق ٧١ الى الفرق اللاحق ١٢٣ كالآتي :

$$\frac{\text{س}}{٧١} = \frac{١٠ - \text{س}}{١٢٣}$$

$$\therefore \frac{\text{س}}{١٠ - \text{س}} = \frac{٧١}{١٢٣}$$

$$\therefore ١٢٣ \text{ س} = ٧١٠ - ٧١ \text{ س}$$

$$\therefore ١٩٤ \text{ س} = ٨١٠$$

$$\therefore \text{س} = ٣٫٦٦$$

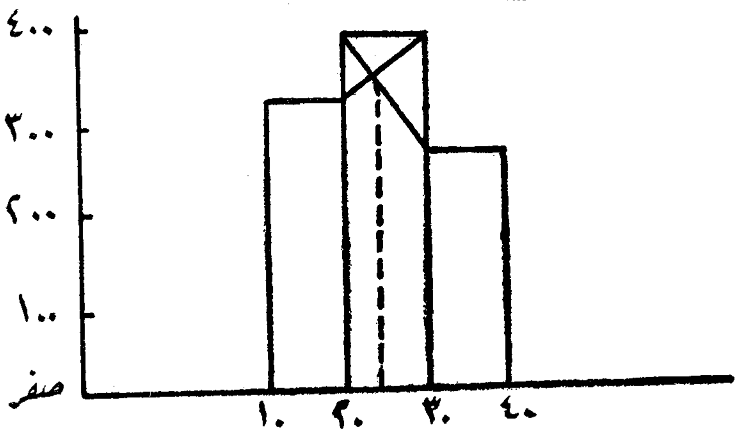
$$\therefore \text{المنوال} = ١٠ + ٣٫٦٦ = ١٣٫٦٦$$

ويمكن وضع هذه الطريقة في صيغة قانون كالآتي :

$$\text{المنوال} = \text{الحد الادنى للفئة المنوالية} + \frac{\text{الفرق السابق}}{\text{الفرق السابق} + \text{الفرق اللاحق}} \times \text{طول الفئة}$$

ونلاحظ ان المنوال محسوباً بالطريقة الاولى يختلف عن قيمته محسوباً بالطريقه الثانية ، وليس هذا من المستغرب حيث ان كلاً من الطريقتين تقريبية كما سبق ان ذكرنا وخاصة اذا تذكرنا ان عدم الدقة يزداد كلما زاد مدى الفئة ، وفي هذا التوزيع مدى الفئة واسع نوعاً ما .

ويمكن التعبير عن هذه الطريقة بيانياً بالشكل الآتي وهو يوضح بجلاء كيف تؤثر الفروق بين التكرارات السابقة وتكرار الفئة المنوالية على موضع المنوال في هذه الفئة ويمكن في الواقع استخدام هذه الطريقة البيانية والاستعاضة بها عن طريقة الفروق الحسابية .



ونلاحظ ان المنوال يقرب من الحد الادنى للفئة المنوالية كلما كان الفرق السابق أصغر من الفرق اللاحق ، والعكس يقرب المنوال من الحد الاعلى للفئة المنوالية كلما كان الفرق اللاحق أصغر من الفرق السابق ، فاذا تساوى الفرقان يكون المنوال هو مركز الفئة المنوالية ، وفي هذه الحالة تكون قيمة المنوال محسوبة بهذه الطريقة مساوية تماماً لقيمته محسوبة بطريقة الرافعة حيث يكون الفرقان متساويان عندما يكون التكرار السابق واللاحق متساويان وبذلك تتوازن الرافعة في منتصفها .

وفي بعض الحالات يكون تحديد المنوال أمراً معقداً لسبب وجود أكثر من قيمة واحدة تتركز عندها القيم . ويحدث ذلك إما لأن الفئات صغيرة المدى جداً بالنسبة الى عدد الوحدات التي نقوم بدراستها وبذلك يكون توزيعها في هذه الفئات لا يساعد على إظهار الفئة التي تتركز فيها القيم . وفي هذه الحالة يمكن التخلص من تعدد قمم المنحنى بتوسيع مدى الفئة شيئاً فشيئاً حتى نحصل على فئة منوالية واحدة ، واذا بقي التوزيع متعدد القمم بالرغم من ذلك فان هذا يكون دليلاً قاطعاً كما سبق ان ذكرنا على عدم تجانس البيانات التي نقوم بدراستها وفي هذه الحالة لا يكون للمنوال أي أهمية تذكر .

الوسط الهندسي :

الوسط الهندسي لأي مجموعة من القيم عددها ن هو الجذر النوني لحاصل ضرب هذه القيم .

مثـال ١٥ :

الوسط الهندسي للقيم ٢ ، ٥ ، ٦ ، ٦ ، ٧ ، ١٠ ، ١٢

$$\sqrt[٧]{٢ \times ٥ \times ٦ \times ٦ \times ٧ \times ١٠ \times ١٢}$$

∴ لو الوسط الهندسي $= \frac{١}{٧}$ (لو ٢ + لو ٥ + لو ٦ + لو ٦ + لو ٧ + لو ١٠ + لو ١٢)

$$= \frac{١}{٧} \times ٥٫٦٥٦٧ = ٠٫٨٠٨١$$

∴ الوسط الهندسي = ٦٫٤٣ الى أقرب رقمين عشريين .

واذا كانت القيم غير متساوية في الأهمية أي ان كلاً منها له وزن معين يدل عليه عدد مرات تكرارها يكون الوسط الهندسي هـو جذر مجموع الاوزان لحاصل ضرب القيم مرفوعة الى أوزانها .

مثـال ١٦ :

القيم	٢	٥	٢٦	٧
الأوزان	٢٥	٣٨	١٩	١٨

∴ الوسط الهندسي $= \sqrt[١٠٠]{٢^{٢٥} \times ٥^{٣٨} \times ٦^{١٩} \times ٧^{١٨}}$

∴ لو الوسط الهندسي = $\frac{١}{١٠٠}$ (٢٥ لو ٢ + ٢٨ لو ٥ + ١٩ لو ٦
+ ٢٨ لو ٧)

= $\frac{١}{١٠٠}$ (٢٥ × ٠ر٣٠١ + ٣٨ × ٦ر٩٩ + ١٩ × ر٧٧٨٢ +
١٨ × ٠ر٨٤٢٧)

= $\frac{١}{١٠٠}$ × ٦٣ر٠٨٤٦

= ر٦٣٠٨٤٦

∴ الوسط الهندسي = ٢٧ر٤

واذا كانت القيم مبوبة في توزيع تكراري نتبع نفس الخطوات في المثال السابق حيث ان القيم المبوبة ليست إلا قيم ذات اوزان مختلفة تدل عليها تكرارات القيم . وتكون الخطوة الاولى في العمل هي ايجاد مراكز الفئات ثم ايجاد لوغاريتمات هذه المراكز ثم ضرب لوغاريتم كل مركز × تكرار المركز وجمع ناتج الضرب . وبقسمةهذا المجموع على مجموع التكرارات تحصل على لوغاريتم الوسط الهندسي ومن جدول الاعداد المقابلة نحصل على القيمة المقابلة لهذا اللوغاريتم تكون هي الوسط الهندسي .

مثـال ١٦ :

ف	ك	م	لو م	لو م × ك
صفر-	٢٦٢	٥	٦٩٨٩٧ر	١٨٣ر١٣٠١٤
١٠-	٣٣٣	١٥	١ر١٧٦٠٩	٣٩١ر٦٣٧٩٧
٢٠-	٢١٠	٢٥	١ر٣٩٧٩٤	٣٦٣ر٥٦٧٤٠
٣٠-	١٤٢	٣٥	١ر٥٤٤٠٧	٢١٩ر٢٥٧٩٤
٤٠-	٨٥	٤٥	١ر٦٥٣٢١	١٤٠ر٥٢٣٨٥
٥٠-	٣٢	٥٥	١ر٧٤٠٣٦	٥٥ر٦٩١٥٢
٦٠-	٢٠	٦٥	١ر٨١٢٩١	٣٦ر٢٥٨٢٠
٧٠-	١٥	٧٥	١ر٨٧٥٠٦	٢٨ر١٢٥٩٠
٨٠-	١٢	٨٥	١ر٩٢٩٤٢	٢٣ر١٥٣٠٤
٩٠-	٨	٩٥	١ر٩٧٧٧٢	١٥ر٨٢١٧٦
١٠٠- ١١٠	٢	١٠٥	١ر١٢١١٩	٤ر٠٤٢٤٨
	١١٢١			١٣٩١ر٢٠٩١٠

$$\therefore \text{لو الوسط الهندسي} = \frac{\text{٢٠٩١ ر ١٣٩١}}{\text{١١٢١}} = \text{١ر٢٤١٠٤}$$

$$\therefore \text{الوسط الهندسي} = \text{١٧ر٤١٨}$$

الوسط التوافقي

الوسط التوافقي لأي عدد من القيم هو مقلوب الوسط الحسابي لمقلوب القيم ، فاذا كانت القيم هي $س_1$ ، $س_2$ ، $س_3$ ..... $س_ن$

$$\therefore \text{الوسط التوافقي} = \frac{ن}{\frac{1}{س_1} + \frac{1}{س_2} + \frac{1}{س_3} \cdots\cdots \frac{1}{س_ن}}$$

وبشكل آخر يكون $\frac{١}{\text{الوسط التوافقي}} = \frac{\text{مجـ}\frac{١}{\text{س}}}{\text{ن}} = \frac{١}{\text{ن}}\ \text{مجـ}\ \frac{١}{\text{س}}$

مثال ١٧ :

الوسط التوافقي للقيم ٢ ، ٤ ، ٨ $= \frac{٣}{\frac{١}{٢} + \frac{١}{٤} + \frac{١}{٨}} = \frac{٣}{\frac{٧}{٨}}$

$= ٣٫٤٣$

واذا كانت القيم مبوبة في توزيع تكراري نجد اولاً مراكز الفئات ثم نحسب مقلوباتها من الجدول الخاص بذلك ، ثم نضرب كل مقلوب في تكرار الفئة ثم نجمع حواصل الضرب ثم نقسم مجموع التكرارات على الناتج تكون النتيجة هي الوسط التوافقي كما هو مبين في المثال التالي :

مثال ١٨ :

ف	ك	م	مقلوبات المراكز	المقلوب × التكرار
صفر–	٢٦٢	٥	٠٫٢	٥٢٫٤
١٠ –	٣٣٣	١٥	٠٫٠٦٦٦	٢٢٫١٧٧٨
٢٠ –	٢١٠	٢٥	٠٫٠٤	٨٫٤
٣٠ –	١٤٢	٣٥	٠٫٠٢٨٦	٤٫٠٦١٢
٤٠ –	٨٥	٤٥	٠٫٠٢٢٢	١٫٨٨٧٠
٥٠ –	٣٢	٥٥	٠٫٠١٨٢	٠٫٥٨٢٤
٦٠ –	٢٠	٦٥	٠٫٠١٥٤	٠٫٣٠٨٠
٧٠ –	١٥	٧٥	٠٫٠١٣٣	٠٫١٩٩٥
٨٠ –	١٢	٨٥	٠٫٠١١٨	٠٫١٤١٦
٩٠ –	٨	٩٥	٠٫٠١٠٥	٠٫٠٨٤٠
١٠٠–١١٠	٢	١٠٥	٠٫٠٠٩٥	٠٫٠١٩٠
	١١٢١			٩٠٫٢٦٠٥

$$\therefore \text{الوسط التوافقي} = \frac{١١٢١}{٩٠,٢٦٠٥} = ١٢,٤١٩$$

ويمكن التعبير عن الخطوات السابقة بالقانون :

$$\text{الوسط التوافقي} = \frac{\text{ن أو مجـ ك}}{\text{مجـ } \frac{\text{ك}}{\text{س}}}$$

ترجيح المتوسطات :

اذا طلب منا ان نحسب متوسط سعر القمح على اختلاف أنواعه واذا علمنا ان سعر الكيلو من أصنافه الثلاث هي ٤٠ ، ٤٣ ، ٤٥ قرشاً فانه من السهل ان نحسب المتوسط المطلوب بجمع هذه الأسعار وقسمة المجموع على عدد الأصناف ويسمى المتوسط في هذه الحالة بسيطاً حيث أننا لم نفرق بين أهمية صنف وآخر وهذا ما يجعل المتوسط الناتج لا يعبر تعبيراً دقيقاً عن متوسط سعر القمح في السوق نظراً لاختلاف أصنافه المختلفة ، ذلك الاختلاف الذي يظهر في كمية المبيع من هذه الأصناف. فاذا فرض ان نسبة ما بيع من الصنف الاول الى ما بيع من الصنف الثاني الى ما بيع من الصنف الثالث هي كنسبة ٣ : ١ : ٦ فمتوسط السعر حينئذ لا بد أن يرجح وفقاً لتلك الأهميات هكذا:

$$\frac{٤٠ \times ٣ + ٤٣ \times ١ + ٤٥ \times ٦}{١٠} = ٤٣,٣ \text{ قرشاً .}$$

والمتوسط بهذه الطريقة لا بد ان يختلف عن المتوسط السابق ولكنه يكون أكثر دقة منه .

ومن الواضح ان كل متوسط يحسب لتوزيع تكراري هو في الواقع متوسط مرجح وأساس الترجيح وبمعنى آخر الأوزان المستخدمة في الترجيح

هي تكرارات الفئات المختلفة ، ذلك لأننا بضرب مركز الفئة × تكرارها كأننا نرجح هذا المركز وهو القيمة بعدد تكرارات هذه القيمة .

ولهذا لو اقتضى الأمر ترجيح الوسط حسابي أو الهندسي أو التوافقي لعدة قيم بسيطة فيتبع لحساب هذه المتوسطات الطريقة التي شرحناها عند الكلام على حساب متوسطات التوزيعات التكرارية .

ومن أهم ميادين استخدام فكرة الترجيح - حساب متوسط ما طرأ على نفقات المعيشة من اختلاف بين تاريخ وآخر فلو علمنا ان أسعار الأغذية ارتفعت في سنة ١٩٤٩ عنها في سنة ١٩٣٩ بمقدار ٣٠٠٪ وان ايجارات المساكن زادت بنسبة ٢٠٠٪ وأسعار الملابس بمقدار ١٥٠٪ عما كانت عليه فلا بد لكي نحسب متوسط ما طرأ على نفقات تلك البنود الثلاثة من تحديد الأهمية النسبية لها ضمن نفقات الشخص العادي وتستخدم تلك الأهمية النسبية أساساً للترجيح (يمكن معرفة الأهمية النسبية لنفقات الشخص العادي على أبواب الأنفاق المختلفة من بحث لميزانية الاسرة يقوم على أساس المعاينة) .

ومما يستخدم فيه مبدأ الترجيح أيضاً - حساب معدلات الوفيات - فلو علمنا أن معدل وفيات الأطفال في بلد ما هو ٧٠٪ ومعدل وفيات الشبان ١٩٪والكهول٥٠٪فانه يستحيل علينا حساب متوسط معدل الوفاة العام دون تحديد الأهمية العددية النسبية للأطفال والشبان والشيوخ في المجتمع موضوع الدراسة لاستخدامها في ترجيح تلك المعدلات .

خصائص المتوسطات المختلفة ومجالات استخدامها :

الوسط الحسابي - يستخدم الوسط الحسابي في أغلب الحالات نظراً لبساطة ووضوح فكرته الأساسية ولشيوع استعماله.وهو يمتاز عن المتوسطات الأخرى من الناحية الفنية لامكان معالجته رياضياً ، الأمر الذي يجعل منه وسيلة قوية في البحث الاحصائي .

ومن أهم تلك الصفات كون المجموع الجبري لانحرافات القيم عنه يساوي صفراً وهذه الخاصية هي الأساس الذي تبني عليه الطريقة المختصرة في حساب هذا المتوسط .

كذلك نلاحظ أن مجموع مربع انحرافات القيم عن الوسط الحسابي أقل ما يمكن أي أقل من مجموع مربع انحرافات القيم عن أي قيمة أخرى غير المتوسط . وسوف تساعدنا هذه الخاصية في قياس التشتت بين القيم .

كذلك نلاحظ أن الوسط الحسابي لمجموعة مكونة من مجموعات صغيرة يساوي الوسط الحسابي المرجح لمترسطات تلك المجموعات الجزئية وهذه صفة لا تتوفر في غيره من المتوسطات .

كما أن الوسط الحسابي لمجموعة من القيم اذا ضرب في عدد هذه القيم فان الناتج يساوي مجموع هذه القيم .

ومن أهم عيوب الوسط الحسابي كونه يتأثر بجميع القيم في المجموعه على حد سواء فاذا تصادف وجود قيم شاذة في المجموعة انحرف الوسط الحسابي تبعاً لذلك . ولما كانت القيم الشاذة في أي توزيع تكراري توجد غالباً في الطرفين لهـــذا يمكننا أن نقول أن الوسط الحسابي يتأثر بالقيم المتطرفة ولا يجب استخدامه في مثل هذه الحالات .

ولا يمكن تقدير الوسط الحسابي لتوزيع تكراري اذا كانت احدى فئاته مفتوحة اذ يلزم العمل الحسابي لتقديره تعيين مراكز الفئات جميعها ولا يمكننا بطبيعة الحال تحديد مركز فئة لم يعين لها أحد حديها (الا في الحالات الـتي تسمح طبيعة البيانات افتراض حد معقول لها) .

ومـــا يجب الاشارة اليه انه تبعاً لاحدى صفات الوسط الحسابي يمكن حسابه دون ضرورة معرفة القيم كلها واحدة واحدة اذ يكفي أن نعرف مجموع هذه القيم وعددها لكي نقدر وسطها الحسابي . فاذا أمكن الحصول على

مجموع ما يدفع لعمال مصنع معين من الأجور وعدد هؤلاء العمال امكن بقسمة الأول على الثاني الحصول على متوسط الاجر في هذا المصنع وهو ما لا يمكن عمله فيما يختص بالمتوسطات الأخرى .

الوسيط – أهم ما يمتاز به الوسيط عن غيره من المتوسطات امكان تقديره بدقة للتوزيعات التكرارية ذات الفئات المفتوحة وكذلك للتوزيعات المختلفة المدى أن التوريعات الغير منتظمة – وكذلك يمكن حسابه بالرسم البياني بدقة .

كما وان الوسيط يمتاز بكونه لا يتأثر بالقيم الشاذة في التوزيعات فهو دائماً يقع بين الوسط الحسابي والمنوال في التوزيعات المعتدلة الالتواء . وهي خاصية سوف نشرحها بالتفصيل عند الكلام على العلاقة بين المتوسطات الثلاث – الوسط الحسابي والمنوال والوسيط .

ونلاحظ أن الوسيط له خاصية هامة وهي أن مجموع الانحرافات المطلقة للاعداد عن وسيطها أصغر من مجموع انحرافاتها المطلقة عن أي عدد آخر (الصناعة التي عند الوسيط الجغرافي) .

ويستعمل الوسيط في التعبير عن متوسط التراتيب حيث يكون من غير المتطق استخدام الوسط الحسابي لأنه يتفق أكثر مع الاعداد التي تعبر عن القياس والـلم . كذلك يستخدم الوسيط في حالة التجارب التي تسبب اتلاف المفردات المختبره أثناء اجراء التجربة مثل تجربة خاصة بمتوسط عمر اللمبة الكهربائية .

المنوال – يمتاز المنوال ببساطة فكرته ووضوحها اذ يمكن فهم معناه كمتوسط دون كبير عناء وقد يكون في هذه الصفة اسبق جميع المتوسطات الأخرى حتى الوسط الحسابي، فمن السهل أن ندرك معنى القيمة الأكثر تكراراً في مجموعة من القيم .

وكذلك يتميز المنوال بأنه لا يتأثر بالقيم الشاذة المتطرفة لأنه غالباً يكون وسط التوزيع فلا يتأثر بالأطراف .

على انه يجب أن نلاحظ ان المنوال لا يكون له أي دلالة إذا لم يكن هناك اتجاه واضح نحو التركز في التوزيع التكراري . وكذلك نلاحظ أنه قيمة تقريبية وقد سبق الاشارة إلى هذه الصفة للمنوال ، ولهذا لا يمكن استخدامه في المسائل التي تقتضي معالجة جبرية، ولذا فانه يختفي في الموضوعات الاحصائية الراقية بعكس الحال في الوسط الحسابي والوسط الهندسي .

ويستحسن عدم استخدام المنوال في الحالات التي يكون التوزيع فيها مفتوحاً فقيمته تعتمد على التكرارات وهذه تتوقف على أطوال الفئات فلا بد من الوقوف على طول الفئة قبل الحكم على كونها منوالية أم لا. ولذلك يستحسن عدم استخدام المنوال في التوزيعات الغير متساوية الفئات حيث أن الطرق التي شرحناها سابقاً تفترض ضمناً التساوي في الفئات .

ويعاب على المنوال انه يتأثر بطريقة اختيار فئات التوزيع أي ان القيمة التي نعتبرها ممثلة للمجموعة تتوقف على طريقة تلخيص الباحث لبياناته وهذه حقيقة تجعل المنوال لا يوثق به كممثل للمجموعة .

ولا يصح استخدام المنوال في حالة التوزيعات التكرارية الملتوية وخصوصاً الحادة الالتواء منها فكما أن الوسط الحسابي ينحرف كثيراً في اتجاه ذيل التوزيع نجد أن المنوال ينحرف الى الجهة الأخرى .

الوسط الهندسي – يخضع الوسط الهندسي للمعالجة الجبرية مثل الوسط الحسابي ، فان حاصل ضرب مجموعة من القيم لا يتغير إذا استبدل كل من هذه القيم بالوسط الهندسي ، فحاصل ضرب القيم ٢ ، ٤ ، ٨ = ٦٤ والوسط الهندسي لهذه القيم يساوي ٤ ، فاذا ضربنا ٤ في نفسها ثلاث مرات فان الناتج يكون ٦٤ .

- يجب وضع الوسط الهندسي بسطاً للقيم الصغيرة وأما الكبيرة ع...
والقيم الصغيره مقاماً للوسط ... للوسط الهن...

وكما انه في حالة الوسط الحسابي يتساوى مجموع انحرافات القيم التي تقل عن الوسط الحسابي مع مجموع انحرافات القيم التي تزيد عنه ، فان حاصل ضرب نسب القيم التي تقل عن الوسط الهندسي إلى هـذا الوسط تساوي حاصل ضرب نسبة القيم التي تزيد عنه فمثلاً القيم ٣ ، ٦ ، ٨ ، ٩ الوسط الهندسي لها = ٦ وبذلك يكون $\frac{٦}{٣} \times \frac{٦}{٦} = \frac{٨}{٦} \times \frac{٩}{٦}$.

وتبعاً لهذه الخاصية يستخدم الوسط الهندسي في حساب متوسطات النسب (الأرقام القياسية على وجه الأخص) وإذا استخدم الوسط الحسابي بدلا من الهندسي لايجاد متوسط النسب فان النتيجة ستكون أعلى من الحقيقة ولهذا يقال ان الوسط الحسابي متحيز إلى أعلى ، فلو فرض مثلاً ان سعر سلعة ما قد زاد إلى الضعف ونقص سعر سلعة اخرى الى النصف فمن المعقول أن نقرر أن مستوى سعر السلعتين معاً لم يتغير ولكن لو استخدمنا الوسط الحسابي لظهر ان مستوى السعر لهاتين السلعتين قد زاد إلى $١ \frac{١}{٤}$ مرة عما كان عليه

$\frac{٢ + \frac{١}{٢}}{٢} = ١ \frac{١}{٤}$ في حين أن الوسط الهندسي لمستوى السعرين هو نفسه المستوى القديم $\sqrt{٢ \times \frac{١}{٢}} = ١$. كذلك يستخدم الوسط الهندسي في الموضوعات التي تستلزم حساب متوسط قيم متزايدة أو متناقصة . فمثلاً يمكن حساب معدل الزيادة في مبلغ معين من المال اودع بفائدة مركبة وذلك باستخدام المعادلة .

المبلغ في نهاية المدة = المبلغ في أول المدة $(١ + ر)^{ن}$

ر هي معدل الفائدة

ن هي المـدة .

مثال ١٩ :

إذا أودع مبلغ ١٠٠٠ جنيه بفائدة مركبة في بنك لأصبح ١٦٠٠ جنيه بعد ١٢ سنة فكم يكون معدل الفائدة .

يظهر من ذلك أن هناك زيادة بمعدل ٦٠ ٪ في ١٢ سنة أي ٥ ٪ في السنة ، هذا إذا استخدمنا الوسط الحسابي ؛ ولكن الواقع أن هذا ليس معدل الفائدة الحقيقية ، أما إذا استخدمنا فكرة الوسط الهندسي يكون معدل الفائدة كالآتي :

$$١٦٠٠ = ١٠٠٠ \, (١ + ر)^{١٢}$$

$$ر = \sqrt[١٢]{\frac{١٦٠٠}{١٠٠٠}} - ١$$

$$ر = ٠٤ ,$$

فإذا أردنا أن نحسب قيمة هذا المبلغ بعد ستة سنوات بنفس الفائدة المركبة نجده يساوي $١٠٠٠ \, (١ + ٠٤ ر)^{٦} = ١٢٦٥$ جنيه — هذه النتيجة يمكن الحصول عليها بايجاد الوسط الهندسي للمبلغين ١٠٠٠ ، ١٦٠٠ فيكون :

$$\sqrt[٢]{١٠٠٠ \times ١٦٠٠} = ١٢٦٥ .$$

مثال ٢٠ :

إذا فرضنا أن عدد السكان في سنة ١٩٢٧ = ٩ مليون
وعدد السكان في سنة ١٩٣٧ = ١٦ مليون

فإذا أردنا ايجاد عدد السكان في سنة ١٩٣٢ فهناك طريقتان :

الطريقة الاولى : عدد السكان في ١٩٣٢ = $\sqrt[٢]{٩ \times ١٦}$ = ١٢ مليون

الطريقة الثانية : ١٦ = ٩ (١ + ر) ^١٠^

لو ١٦ = لو ٩ + ١٠ لو (١ + ر)

١ر٢٠٤١٢٠ = ر٩٥٥٣٤٣ + ١٠ لو (١ + ر)

٠ر٢٤٨٧٧٧ = ١٠ لو (١ + ر)

١ + ر = ١ر٠٥٩١

ر = ر٠٥٩١

وبذلك يكون معدل زيادة السكان = ١ر٥٩ في الألف .

عدد السكان في سنة ١٩٣٢ = ٩ (١ + ر٠٥٩١)^٥^

= ١٢ مليون .

(استخدمنا فكرة الوسط الهندسي لأن الزيادة في عدد السكان تشبه تماماً الزيادة في مبلغ من المال بفائدة مركبة) .

وأهم ما يعاب على الوسط الهندسي صعوبته الحسابية وضرورة استخدام جداول اللوغاريثمات في ايجاده .

الوسط التوافقي – يقتصر استخدامه في ايجاد متوسطات الأسعار اذا اعطيت بدلالة وحدة النقود ، كذلك في حالة ايجاد متوسط السرعة اذ انها أيضاً تعطي في العادة بدلالة وحدة الزمن فيقال مثلاً (خمسة وحدات من سلعة ما بقرش – أو قطعت سيارة المسافة بسرعة ٩٠ كيلومتراً في الساعة وهكذا) . لو استخدم الوسط الحسابي في أمثال تلك الحالات فانه يكون متحيزا الى أعلى ، فلو فرضنا اننا اشترينا عدداً من الليمون بسعر ثلاثة بقرش ونفس العدد بسعر خمسة بقرش فمن المعقول ان نقرر ن متوسط سعر ما اشتريناه هو أربعة بقرش . نحصل على هذه النتيجة اذا استخدمنا الوسط التوافقي ولكن نحصل على سعر أعلى لو استخدمنا الوسط الحسابي هكذا :

السعر بالقرش هما $\frac{١}{٣}$ ، $\frac{١}{٥}$ ووسطهما الحسابي هو $(\frac{١}{٣} + \frac{١}{٥}) \div ٢ = \frac{٨}{٣٠}$

وهذا العدد أكبر من $\frac{١}{٤}$ وهو الجواب الصحيح . والوسط التوافقي هو

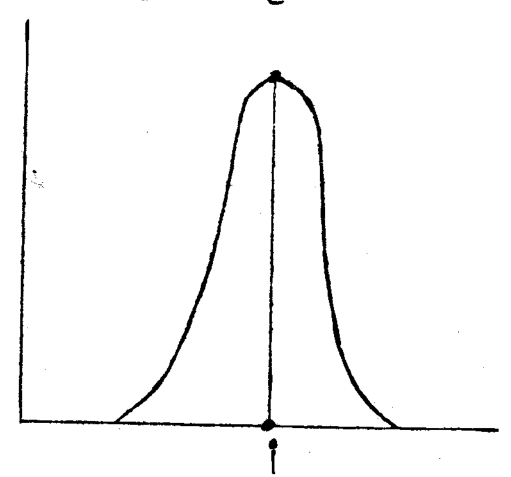
$$\frac{٢}{٣ + ٥} = \frac{١}{٤}$$

وعيب الوسط التوافقي عدم شيوع استخدامه وغرابته وصعوبة عملياته الحسابية اذ يقتضي استخدام جداول مقلوبات الأعداد وهذه قل ان تتوفر للباحث العادي .

العلاقة بين المتوسطات :

تتـــوقف العلاقة بين المتوسطات الثلاث – الوسط الحسابي والوسيط والمنوال – على نوع التوزيع التكراري وبالتالي على المنحنى الذي يمثل هذا التوزيع .

فإذا كان التوزيع تام التماثل يمثله منحنى معتدل طبيعي فان المتوسطات الثلاث السابقة تتساوى تماماً أي تنطبق على بعضها عند القيمة المقابلة لقمة المنحنى التكراري الذي يمثل التوزيع كما يظهر في الشكل الآتي .



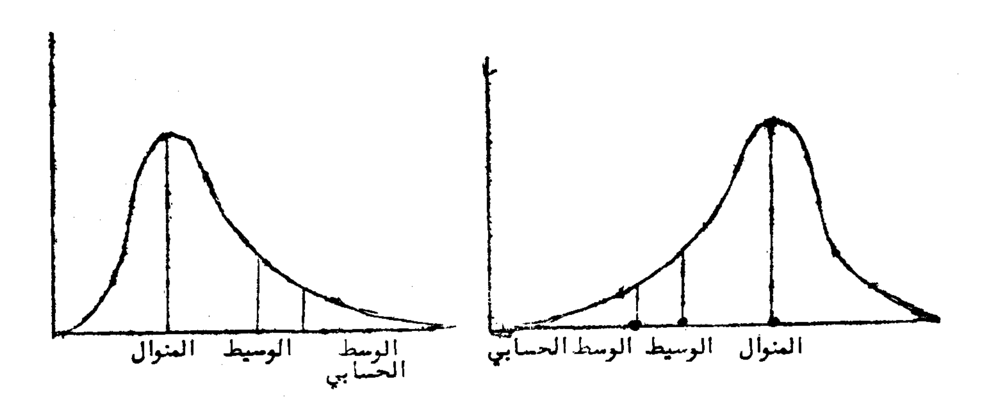
(النقطة أ تمثل الوسط الحسابي والوسيط والمنوال)

الوسط الحسابي - المنوال = ٣ (الوسط - الوسيط)

وإذا كان التوزيع معتدل الالتواء فإن الوسيط يقع دائماً بين المتوسطين الآخرين والمسافة بينه وبين الوسط الحسابي تساوي ثلث المسافة بين الوسط الحسابي والمنوال ، وهذه العلاقة يمكن وضعها في صيغة المعادلة الآتية :

الوسط الحسابي — المنوال = ٣ (الوسط — الوسيط)

وتتضح هذه العلامة من الرسم البياني الآتي :



تمـــارين

١ — اذا كان متوسطات الدرجـات في احدى المواد لثلاث فصول هي ٧٥ ، ٨٢ ، ٨٤ وكان عدد التلاميذ في هذه الفصول ، ٣٢ ، ٢٥ ، ١٧ — احسب المتوسط العام لدرجات الثلاث فصول : الوسط = ٧٨ .

٢ — اذا كان المتوسط العام لاجور العمال = ٥٠٠٠ ليرة في العام وكان متوسط الاجور للعمال الذكور = ٥٢٠٠ ليرة ومتوسط الأجور للعمـال الاناث = ٤٢٠٠ ليرة احسب نسبة العمال الذكور والاناث — النسبة هي ٨٠٪ ، ٢٠٪ .

٣—التوزيع الآتي يبين قوة احتمال الكابلات التيأنتجتها احدى المؤسسات.

قوة الاحتمال (طن)	عدد الكابلات
٩ر٣ - ٩ر٧	٢
٩ر٨ - ١٠ر٢	٥ر
١٠ر٢ - ١٠ر٧	١٢
١٠ر٨ - ١١ر٢	١٧
١١ر٣ - ١١ر٧	١٤
١١ر٨ - ١٢ر٢	٦
١٢ر٣ - ١٢ر٧	٣
١٢ر٨ - ١٣ر٢	١

احسب متوسط قوة احتمال الكابل الواحد بالطريقة المطولة والطريقة المختصرة :

س- = ١١ر٠٩ طن

٤ - الآتي التوزيع النسبي للمبات الكهربائية حسب طول عمرها :

عمر اللمبة بالساعة	٪ من مجموع اللمبات
١٠٠ -	٥ر٦
٢٠٠ -	١٤ر٤
٣٠٠ -	١٤ر٦
٤٠٠ -	٢٥ر٤
٥٠٠ -	١٩ر٤
٦٠٠ -	١٦ر٦
٧٠٠ - ٨٠٠	٤ر٠

احسب الوسط الحسابي والوسط الهندسي والوسط التوافقي .

٥ – الآتي توزيع أعمار أرباب الاسر :

العمر بالسنوات الكاملة	عدد الاسر بالألف
أقل من ٢٥	٢
٢٥ – ٢٩	٤
٣٠ – ٣٤	٥
٣٥ – ٤٤	١٠
٤٥ – ٥٤	٩
٥٥ – ٦٤	٦
٦٥ – ٧٤	٤
٧٥ وأكثر	١

احسب الوسيط والربيعين (حسابياً وبالرسم البياني)

٦ – احسب المنوال للتوزيع الآتي (بالطرق المختلفة)

الفئـــات	التكرار
١٠ –	٣
١٥ –	٧
٢٠ –	١٦
٢٥ –	١٢
٣٠ –	٩
٣٥ –	٥
٤٠ – ٤٥	٢

المنوال = ٥ر٢٣

الفصل السادس

التشتت

كان اهتمامنا في الفصول السابقة موجهاً الى كيفية تلخيص عدد كبير من البيانات الرقمية تلخيصاً يساعد على فهمها وتحليلها احصائياً ، ثم الى محاولة البحث عن قيمة واحدة يمكن ان توصف بها هذه البيانات وهي تلك القيمة التي نعتقد ان القيم الاخرى تميل الى التركز حولها ولكن أي متوسط لا يكفي وحده لقياس هذا الاتجاه نحو التركز ، ولذلك نكون في حاجة الى ثلاث أنواع أخرى من المقاييس حتى نستطيع ان نصف التوزيع التكراري وصفاً كاملا وحتى يمكن مقارنته بتوزيع تكراري آخر . هذه المقاييس الثلاث هي أولاً مقياس يقيس لنا مدى تباعد أو تركز القيم حول القيمة المتوسطة وبمعنى آخر مدى بعثرة القيم أو اختلافها أو تشتتها ، والمقياس الثاني يقيس لنا درجة تماثل القيم حول القيمة المتوسطة والمقياس الثالث يقيس لنا درجة تفرطح التوزيع وبمعنى آخر درجة تجمع القيم عند القيمة المنوالية .

أهمية قياس التشتت

التشتت صفة هامة من صفات أي مجموعة من البيانات الرقمية فلا يمكن أن نتصور تساوي أطوال جميع الطلبة أو أوزانهم أو أعمارهم ولا يمكن ان نتصور تساوي الانتاج في جميع المؤسسات الصناعية أو تساوي دخول جميع أفراد المجتمع ، ولهذا فان القيمة التي نعتبرها ممثلة لمجموعة من القيم لا بد وان

تكون مصحوبة بقيمة أخرى تقيس لنا مدى تباعد القيم أو قربها من هذه القيمة ، بحيث اذا كبر هذا المقياس الى درجة بعيدة فان المتوسط يفقد أهميته كقيمة ممثلة ، أما اذا صغر مقياس التشتت تزداد أهمية المتوسط كقيمة ممثلة ، فيمكننا الاعتماد عليه في بحثنا لهذه البيانات بحثاً احصائياً

مقاييس التشتت

يمكن قياس التشتت بمقاييس مميزة بنفس وحدات البيانات التي نقوم بدراستها أو بمقاييس مطلقة أي كنسبة مئوية مستقلة عن الوحدة التي تقاس بها البيانات . فاذا استعملنا مقياساً من النوع الأول فاننا بذلك نقيس التشتت المطلق ، واذا استعملنا مقياساً من النوع الثاني فاننا نقيس التشتت النسبي . وبالطبع تتميز مقاييس النوع الثاني بأنها صالحة للمقارنة بين مجموعات مختلفة بينما لا يمكن استخدام مقاييس النوع الأول في المقارنة .

وأبسط مقاييس التشتت هو المدى أي مدى تغير الظاهرة موضوع الدراسة. فلو أعطينا أجر مائة عامل مثلاً فأبسط مقياس للتشتت بين أجور هؤلاء العمال هو الفرق بين أكبر وأصغر أجر فكلما زاد المدى أو الفرق بين هاتين القيمتين زاد عدم التجانس أي زاد تشتت المجموعة والعكس كلما نقص المدى .

وعيب هذا المقياس عدم دقته حيث انه يتوقف على قيمتين فقط فلو تصادف ان كانت قيم شاذة أصبح المدى لا يعبر تعبيراً دقيقاً عن التشتت بين البيانات . ولهذا فبالرغم من بساطة هذا المقياس الا أننا لا نستطيع أن نعتمد عليه في قياس التشتت أو في مقارنته بين مجموعات مختلفة وهذا هو السبب في أننا نلجأ الى مقاييس أخرى للتعبير عن التشتت لنتخلص بواسطتها من أثر القيم المتطرفة التي تكون أحياناً واضحة الشذوذ .

المدى الربيعي = الربيع الأعلى - الربيع الأدنى
لنصف المدى الربيعي = ½ (الربيع الأعلى - الربيع الأدنى)

نصف المدى الربعي

يعالج عيب المدى الذي أشرنا إليه بقياس التشتت بمقياس لا يتأثر بالقيم المتطرفة وذلك بأخذ الفرق بين الربع الأعلى والأدنى . وحيث اننا نعلم ان نصف عدد القيم ينحصر بين الربيعين وان فيهذا النصف يقع المتوسط الحسابي والوسيط والمنوال وهي القيم التي اعتبرناها ممثلة للتوزيع ، لهـــذا يكون من المنطق أن نعتمد الى حد ما على قياس التشتت بواسطة نصف الفرق بـــين الربيع الأعلى والادنى .

ولحساب هذا المقياس نرتب القيم ترتيباً تصاعدياً أو تنازلياً ثم نحسب قيمة كل منالربيع الأدنى والأعلى بالطريقةالتي شرحناها فيالفصل السابقويكون :

المدى الربيعي = الربيع الاعلى – الربيع الادنى .

نصف المدى الربيعي = $\frac{١}{٢}$ (الربيع الاعلى – الربيع الأدنى) .

التشتت حول المتوسط

في قياس التشتت بالطريقتين السابقتين كنا نقيس مدى الاختلاف بـين القيم المختلفة التي تتخذها ظاهرة ما في تغيرها . ولكننا قد اتفقنا على أن نمثل هذه القيم بقيمة واحدة هي المتوسط ولذلك يكون من المنطق ان نحاول قياس تشتت القيم حول هذا المتوسط . وقد أشرنا في الفصل السابق الى ان الوسط الحسابي يمكن ان يخضع للمعالجة الجبرية ولهـــذا نفضل قياس التشتت حوله .

الانحراف المتوسط :

هو متوسط مجموع انحرافات القيم عن الوسط الحسابي ، ولكننا نعرف أن المجموع الجبري لانحرافات القيم حول الوسط الحسابي يساوي صفراً ، لهـذا

الانحراف المتوسط = هو متوسط مجموع انحرافات القيم عن الوسط الحسابي .

يجب عند حساب الانحراف المتوسط أن نتخلص من الاشارات الجبرية بوسيلة ما ، وأبسط وسيلة لذلك هي مجرد اهمال هذه الاشارات .

وطريقة حساب الانحراف المتوسط في المجموعات الصغيرة (القيم البسيطة الغير مبوبة) هي ايجاد الوسط الحسابي للقيم ثم ايجاد انحرافات هذه القيم عن الوسط الحسابي ، ثم إيجاد مجموع هذه الانحرافات بعد اهمال اشاراتهـــا ، وبقسمة هذا المجموع على عدد القيم ينتج الانحراف المتوسط .

فمثلاً في المجموعة ٣ ، ٦ ، ٩ ، ١٢ ، ١٥ ، المتوسط = ٩

والانحرافات بدون اشارة هى ٦ ، ٣ ، صفر ، ٣ ، ٦ فيكون الانحراف المتوسط = $\frac{١٨}{٥}$ = ٣٫٦ .

أما في حالة التوزيع التكراري فتكون خطوات العمل كالآتي اذا أردنا أن نحسب الانحرافات عن الوسط الحسابي رأساً .

١ – حساب الوسط الحسابي .

٢ – حساب الانحرافات (اهمال الاشارة الجبرية) عن الوسط الحسابي .

٣ – ايجاد حاصل ضرب تكرار كل فئة في انحراف مركز الفئة عن الوسط الحسابي بدون اشارة .

٤ – جمع هذه الانحرافات وقسمة الناتج على التكرار الكلي فينتج لدينا الانحراف المتوسط .

وتتضح هذه الخطوات من الجدول الآتي :

مثال ٢١ :

الفئات	التكرار	المراكز	م × ك	الانحرافات	ح × ك
٠ — ٠	٢٠	٥	١٠٠	١٧	٣٤٠
—١٠	٨٠	١٥	١٢٠٠	٧	٥٦٠
—٢٠	٥٠	٢٥	١٢٥٠	٣	١٥٠
—٣٠	٤٠	٣٥	١٤٠٠	١٣	٥٢٠
—٤٠	١٠	٤٥	٤٥٠	٢٣	٢٣٠
	٢٠٠		٤٤٠٠		١٨٠٠

الوسط الحسابي = $\frac{٤٤٠٠}{٢٠٠}$ = ٢٢

الانحراف المتوسط = $\frac{١٨٠٠}{٢٠٠}$ = ٩ (حسبت الانحرافات عن ٢٢)

ولكن عند حساب الانحرافات عن الوسط الحسابي قد توجد كسور كبيرة في بعض الحالات ، الأمر الذي يؤدي الى تعقيد العمليات الحسابية ، ولهذا نلجأ الى طريقة اخرى لتبسيط العمليات الحسابيـــة لتفادى الكسور في الانحرافات وذلك باستخدام وسط فرضي وتطبيق علاقة جبرية خاصة :

$ح_س$ مجموع الانحرافات عن الوسط الحسابي .

$ح_و$ » » » » الفرضي .

ف الوسط الحسابي — الوسط الفرضي .

$ك_١$ مجموع التكرارات التي تقل عن الوسط الحسابي .

$ك_٢$ » » » تزيد » » »

وبذلك تكون $ح_س = ح_و + ف (ك_١ - ك_٢)$

$ح_س = ح_و + ف (ك_١ - ك_٢)$

والجدول التالي يوضح كيفية استخدام هذه العلاقة في حساب الانحراف المتوسط :

مراكز الفئات	التكرار	ح عن وسط فرضي ٢٥	ح بدون اشارة	ح × ك
٥	٢٠	٢٠ −	٢٠	٤٠٠
١٥	٨٠	١٠ −	١٠	٨٠٠
٢٥	٥٠	صفر	صفر	صفر
٣٥	٤٠	١٠ +	١٠	٤٠٠
٤٥	١٠	٢٠ +	٢٠	٢٠٠
	٢٠٠			١٨٠٠

ف = ٢٢ − ٢٥ = − ٣

ك١ = ١٠٠

ك٢ = ١٠٠

ح س = ١٨٠٠ + (− ٣) (١٠٠ − ١٠٠)

= ١٨٠٠

الانحراف المتوسط = $\frac{١٨٠٠}{٢٠٠}$ = ٩

وتفسير هذه الطريقة ان الرقم ١٨٠٠ هو عبارة عن مجموع انحرافات القيم عن وسط فرضي (٢٥) ، ولكننا نريد مجموع انحرافات القيم عن الوسط الحسابي ولهذا نكون في حاجة الى تصحيح الرقم ١٨٠٠ ، وللتصحيح نلاحظ اننا حسبنا انحرافات القيم التي تقل عن الوسط الحسابي (٢٢) عن وسط فرضي (٢٥) ولذلك لا بد من طرح ١٠٠ × (٢٢ − ٢٥) − أما القيم التي تزيد عن الوسط الحسابي فقد حسبنا انحرافاتها عن الوسط الفرضي ٢٥ ولذلك يجب ان نضيف ١٠٠ (٢٢ − ٢٥) وبذلك يجب تصحيح مجموع الانحرافات عن الوسط الفرضي بطرح − ٣ × ١٠٠ وجمع + ٣ × ١٠٠ .

الانحراف المعياري (Standard Deviation)

وهو المقياس الثاني الذي يقيس تشتت القيم حول الوسط الحسابي. والفكرة الأساسية في هذا المقياس هو انه بدلا من إهمال الاشارات الجبرية عند حساب الانحراف المتوسط وهو اجراء غير منطقي من الناحية الجبرية ، لهذا نحاول التخلص من هذه الاشارات بطريقة أخرى أكثر صلاحية وذلك بتربيع الانحرافات فيتحول السالب منها والموجب الى قيم موجبة .

ونلاحظ ان الانحراف المعياري هو أقوى مقاييس التشتت حساسية وأكثرها شيوعاً فتكاد جميع وسائل التحليل الاحصائي تعتمد عليه . ويمكن تعريفه بأنه الجذر التربيعي لمتوسط مجموع مربعات انحرافات القيم عن الوسط الحسابي . فإذا أخذنا القيم ٣ ، ٧ ، ٨ منه فان انحرافاتها عن الوسط الحسابي هي ٣ ، ١ ، ٢ على التوالي ومربعات تلك الانحرافات هي ٩ ، ١ ، ٤

ومتوسط مجموع هذه الانحرافات هو $\frac{٩ + ١ + ٤}{٣}$ = ٤ر٦٧ وجذره التربيعي ٢ر١٦ .

ويكتفي في بعض الأحيان عند قياس التشتت بالقيمة السابقة قبل حساب الجذر التربيعي لها أي ٤ر٦٧ وفي هذه الحالة يطلق على المقياس اسم التباين .

وواضح ان التباين ما هو الا مربع الانحراف المعياري .

ولا يعاب على الانحراف المعياري مقياس للتشتت سوى ان تمييزه من تمييز المتغير الأصلي ولهذا لا يمكن استخدامه في مقارنة التشتت بين مجموعتين لهما تمييز مختلف ، فلو كان لدينا ظاهرتين لمجموعة واحدة كالأجور وعدد أبناء طائفة من العمال وأردنا مقارنة تشتت الأجور بتشتت عدد الأبناء لما أمكننا ذلك باستخدام الانحراف المعياري اذ ان الانحراف المعياري للاجور سيكون ٥ غروش مثلاً بينما الانحراف المعياري لعدد الأبناء سيكون فردين مثلاً ولا يمكننا بطبيعة الحال مقارنة هذا بذاك .

والعيب الثاني للانحراف المعياري تأثره بالوسط الحسابي للمجموعة ولهذا لا يمكن استخدامه في المقارنة حتى بين مجموعتين لهما نفس التميـيز اذا اختلف وسطهما الحسابي ، فلو أخذنا عدداً من طلبة المدارس الابتدائية ومتوسط أعمـــارهم ٩ سنوات وحسبنا الانحراف المعياري للأعمار وكان سنتين مثلاً ثم أخذنا مجموعة اخرى من طلبة الجامعة ومتوسط أعمـــارهم ٢٠ سنة مثلا ووجدنا ان الانحراف المعياري لأعمارهم هو ٥ سنوات مثلاً فلا يمكننا القول بأن تشتت المجموعة الأولى أقل من تشتت المجموعة الثانية حيث ان هذين المقياسين يتأثران بقيمة الوسط الحسابي في كل حالة . ولهذا لمقارنة التشتت بين مجموعتين يستحسن عــــدم استعمال الانحراف المعياري . وقد اشرنا الى كيفية حساب الانحراف المعياري لمجموعة غير مبوبة من القيم ، أمـــا بالنسبة لتوزيع تكراري يكون من المناسب استعمال وسط فرضي بدلاً من الوسط الحسابي الحقيقي وذلك لتسهيل العمليات الحسابية ثم استعمال معادلة تصحيح. وتكون خطوات العمل كما يلي :

١ – نحدد مراكز الفئات ونأخذ أحد هذه المراكز وسطاً فرضياً .

٢ – نحسب انحرافات المراكز الأخرى عن هذا الوسط الفرضي – سيكون بعضها سالباً وبعضها موجباً .

٣ – تضرب تكرار كل فئة في الانحراف المقابل وتجمع نتائج الضرب جمعاً جبرياً .

٤ – تضرب حاصل الضرب مرة اخرى في الانحراف فتنتـــج مجموعة أعداد كلها موجبة وتجمع هذه الأعداد.

٥ – نستخرج الانحراف المعياري باستخدام القانون .

$$\text{الانحراف المعياري ع} \quad (٥) = \sqrt{\frac{\text{مجـ ح}^{٢}\text{ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right)^{٢}}$$

والجدول التالي يوضح عملياً طريقة حساب الانحراف المعياري :

مثال ٢٢ :

الفئات	التكرارات	المراكز	ح عن وسط فرضي ١٠	ح×ك	ح^{٢}×ك
صفر–	٢٠	٥	−٢	−٤٠	٨٠
١٠ –	٨٠	١٥	−١	−٨٠	٧٠
٢٠ –	٥٠	٢٥	صفر	صفر	صفر
٣٠ –	٤٠	٣٥	+١	+٤٠	٤٠
٤٠ –	١٠	٤٥	+٢	+٢٠	٤٠
	٢٠٠			−٦٠	٢٤٠

$$\text{الانحراف المعياري} = ١٠\sqrt{\frac{٢٤٠}{٢٠٠} - \left(\frac{٦٠}{٢٠٠}\right)^{٢}}$$

$$= ١٠\sqrt{١ر٢ - ٠٩ر}$$

$$= ١٠\sqrt{١١ر١} = ١٠ر٥٥$$

نلاحظ اننا اختصرنا الانحرافات على ١٠ لأن الفئات كلها متساويةالمدى ، والمـــدى = ١٠ ولسبب الاختصار نضرب نتيجة الجذر التربيعي في ١٠ لنرجع الأرقام الى أصلها . على ان هذا الاجراء لا يجب أن نتبعه إلا في حالة التوزيع التكراري المنتظم أي التوزيع الذي يتساوى فيه مدى الفئات وقد أشرنا الى هذه الملاحظة في ايجاد الوسط الحسابي .

على ان الجواب الذي نحصل عليه باتباع هذه الطريقة المختصرة التي تقوم على أساس ايجاد الانحرافات عن وسط فرضي هــو نفس الجواب الذي يمكن

ان نحصل عليه لو اتبعنا الطريقة القائمة على أساس ايجاد الانحرافات عن الوسط الحسابي . وليس اتباعنا لهذه الطريقة المختصرة إلا من قبيل تسهيل العمل الحسابي فقط . فالقاعدة العامة للانحراف المعياري هي الجذر التربيعي لمتوسط مجموع مربع انحرافات القيم عن الوسط الحسابي ، ولكن في حالة ايجاد الانحرافات عن الوسط الفرضي لا بد من تصحيح هذه الانحرافات .

مناقشة قوانين الانحراف المعياري

لاحظنا ان القاعدة الأساسية للانحراف المعياري هي :

$$ع = \sqrt{\frac{مج\,(س - \bar{س})^{٢}}{ن}} \qquad (١)$$

نستطيع ان نصع هذا القانون في شكل آخر بتحليل مج $(س - \bar{س})^{٢}$.
نفترض أن لدينا القيم $س_{١}$ ، $س_{٢}$ ، $س_{٣}$ ، $س_{٤}$ ، ووسطها الحسابي $= \bar{س}$

∴ انحراف القيمة الاولى $= س_{١} - \bar{س}$ مربع هذا الانحراف $= س_{١}^{٢} - ٢\,س_{١}\,\bar{س} + \bar{س}^{٢}$

∴ انحراف القيمة الثانية $= س_{٢} - \bar{س}$ ومربع هذا الانحراف $= س_{٢}^{٢} - ٢\,س_{٢}\,\bar{س} + \bar{س}^{٢}$

وانحراف القيمة الثالثة $= س_{٣} - \bar{س}$ ومربع هذا الانحراف $= س_{٣}^{٢} - ٢\,س_{٣}\,\bar{س} + \bar{س}^{٢}$

وانحراف القيمة الرابعة $= س_٤ - \bar{س}$

مربع هذا الانحراف $= س_٤^٢ - ٢س_٤ \bar{س} + \bar{س}^٢$

∴ مجموع مربع انحرافات القيم $= مج س^٢ - ٢ مج س \bar{س} + ن \bar{س}^٢$

∴ مجموع مربع انحرافات القيم $= مج س^٢ - ٢ ن \bar{س}^٢ + ن\bar{س}^٢$

$$∴ مج (س - \bar{س})^٢ = مج س^٢ - ن \bar{س}^٢$$

$$∴ ع = \sqrt{\frac{مج س^٢}{ن} - \frac{ن \bar{س}^٢}{ن}} = \sqrt{\frac{مج س^٢}{ن} - \bar{س}^٢} \quad (٢)$$

وإذا أخذنا الانحرافات عن الوسط الفرضي ويكون القانون :

$$ع = \sqrt{\frac{مج (س - و)^٢}{ن} - (\bar{س} - و)^٢}$$

حيث ان $مج (س - و)^٢$ هي مجموع مربع الانحرافـــات عن الوسط فرضي ولذا يجب تصحيحها حتى تساوي مجموع مربع الانحرافات عن الوسط لحسابي ولهذا أضفنا الجزء $(\bar{س} - و)^٢$. على أن هذا القانون هو مجرد شكل جبري آخر للقوانين السابقة ولاثبات ذلك نفترض أن لدينا القيم :

$س_١ ، س_٢ ، س_٣ ، س_٤$ ووسطها الفرضي $= و$

∴ انحراف القيمة الاولى عن و $= س_١ - و$

ومربع هذا الانحراف $= س_١^٢ - ٢ و س_١ + و^٢$

وانحراف القيمة الثانية عن و = $س_٢$ − و

ومربع هذا الانحراف = $س_٢^٢ - ٢ و س_٢ + و^٢$

وانحراف القيمة الثالثة عن و = $س_٣$ − و

ومربع هذا الانحراف = $س_٣^٢ - ٢ و س_٣ + و^٢$

وانحراف القيمة الرابعة عن و = $س_٤$ − و

ومربع هذا الانحراف = $س_٤^٢ - ٢ و س_٤ + و^٢$

∴ مجموع مربع هذه الانحرافات = $مجـ س^٢ - ٢و مجـ س + ن و^٢$ أي

$مجـ (س - و)^٢ = مجـ س^٢ - ٢ و ن \bar{س} + ن و^٢$

$$\therefore \frac{مجـ (س - و)^٢}{ن} = \frac{مجـ س^٢}{ن} - ٢ و \bar{س} + و^٢$$

$$\therefore ع = \sqrt{\frac{مجـ س^٢}{ن} - ٢ و \bar{س} + و^٢ - \bar{س}^٢ + ٢ و \bar{س} - و^٢}$$

$$\therefore ع = \sqrt{\frac{مجـ س^٢}{ن} - \bar{س}^٢}$$ وهو نفسه القانون الثاني .

وبذلك يكون القانون الذي نستخدمه عند استخدام الوسط الفرضي هو:

$$ع = \sqrt{\frac{مجـ (س - و)^٢}{ن} - (\bar{س} - و)^٢} \quad (٣)$$

هذا القانون هو نفسه الذي نستخدمه في ايجاد الانحراف المعياري للتوزيع التكراري حيث أن مج (س – و)^{٢} يعبر عنها في التوزيع التكراري بالرز مجح^{٢}ك (ح^{٢} هي مربع الانحراف عن الوسط الفرضي وضربناها في ك لان القيم متكررة ولذلك فانحرافاتها ومربع انحرافاتها تتكرر معهـــا) . أما ن فترمز الى مجموع التكرارات في التوزيع التكراري . أما ($\bar{س}$–و)^{٢} فهي مربع الفرق بين الوسط الحسابي والوسط الفرضي وحسب قانون الوسط الحسابي للتوزيع التكراري $\bar{س} = و + \frac{مج ح ك}{مج ك}$

$$\therefore \bar{س} - و = \frac{مج ح ك}{مج ك}$$

$$\therefore (\bar{س} - و)^{٢} = \left(\frac{مج ح ك}{مج ك}\right)^{٢}$$

وبذلك يكون قانون الانحراف المعياري للتوزيع التكراري هو :

$$ع = \sqrt{\frac{مج ح^{٢} ك}{مج ك} - \left(\frac{مج ح ك}{مج ك}\right)^{٢}}$$

ولا يشتمل هذا القانون على الضرب في طول الفئة حيث اننا أشرنا قبل ذلك الى أننا نضرب في طول الفئة عندما نختصر الانحرافات فقط ولذلك فالضرب في طول الفئة ليس شيئاً أساسياً في القانون .

مثال ٢٣ :

س	س − س̄	(س − س̄)٢	س٢	س − و	(س−و)٢
٣	− ٢	٤	٩	− ١	١
٤	− ١	١	١٦	صفر	صفر
٨	+ ٣	٩	٦٤	+ ٤	١٦
	صفر	١٤	٨٩		١٧

و = ٤

∴ ع بالقانون الاول = $\sqrt{\frac{\text{مج}\,(\text{س} - \bar{\text{س}})^{٢}}{\text{ن}}} = \sqrt{\frac{١٤}{٣}} = \text{٢٫١٦}$

∴ ع وبالقانون الثاني = $\sqrt{\frac{\text{مج س}^{٢}}{\text{ن}} - \bar{\text{س}}^{٢}} = \sqrt{\frac{٨٩}{٣} - ٢٥} = \text{٢٫١٦}$

و ع بالقانون الثالث =

$\sqrt{\frac{\text{مج}\,(\text{س} - \text{و})^{٢}}{\text{ن}} - (\bar{\text{س}} - \text{و})^{٢}} = \sqrt{\frac{١٧}{٣} - (١)^{٢}} = \text{٢٫١٦}$

أي ان القوانين الثلاث تعطينا نفس النتيجة وليس ذلك بمستغرب فهذه القوانين ليست إلا صوراً جبرية مختلفة مشتقة من القانون الاول .

معامل الاختلاف :

أشرنا فيما سبق إلى أن الانحراف المعياري بالرغم من دقته وكثرة استعماله

إلا انه لا يفيدنا في حالة مقارنة تشتت مجموعتين احصائيتين لأسباب شرحناها فيما تقدم . ولكي نتمكن من المقارنة يجب الاستغناء عن الوحدات المميزة باستخدام اعداد مجردة خالية من التمييز ، ونحصل على مثل هذه الاعداد بقسمة عددين من نوع واحد أي مميزين بنفس الوحدات . وأكثر مقاييس المقارنة استخداماً هو معامل الاختلاف ونحصل عليه بقسمة الانحراف المعياري للتوزيع على الوسط الحسابي له وضرب الناتج في ١٠٠ . وبمعنى آخر ليس معامل الاختلاف إلا النسبة المئوية للانحراف المعياري منسوبا إلى الوسط الحسابي – وبالطبع كلما كبر معامل الاختلاف كلما دل ذلك على قوة التشتت بين مفردات التوزيع والعكس كلما صغر دل ذلك على ضعف التشتت

ولكن الانحراف المعياري بالرغم من صلاحيته وأهميته لا يمكن حسابه الا في التوزيعات التي نستطيع تحديد وسطها الحسابي وذلك يستحيل في بعض التوزيعات (التوزيعات المفتوحة) ، ولهذا نستعيض عن الانحراف المعياري بنصف المدى الربيعي (يسمى أحياناً الانحراف الربيعي) . ولكن يعاب على هذا المقياس انه مميز بنفس وحدات التوزيع ولذا لا يمكن استخدامه في المقارنة . ويمكن علاج هذا العيب كما فعلنا بالنسبة للانحراف المعياري بنسبته الى الوسيط . وبذلك يكون معامل الاختلاف الربيعي =

$$\frac{\text{نصف المدى الربيعي}}{\text{الوسيط}} \times ١٠٠$$

ولما كان الوسيط في معظم التوزيعات القريبة من التماثل يساوي نصف

مجموع الربيعين فيكون معامل الاختلاف الربيعي =

$$\frac{\frac{\text{الربيع الأعلى} - \text{الربيع الآدنى}}{٢}}{\frac{\text{الربيع الاعلى} + \text{الربيع الادنى}}{٢}} \times ١٠٠$$

وبذلك يكون معامـــل الاختلاف الربيعي =

$$\frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{\text{الربيع الاعلى} + \text{الربيع الادنى}} \times ١٠٠$$

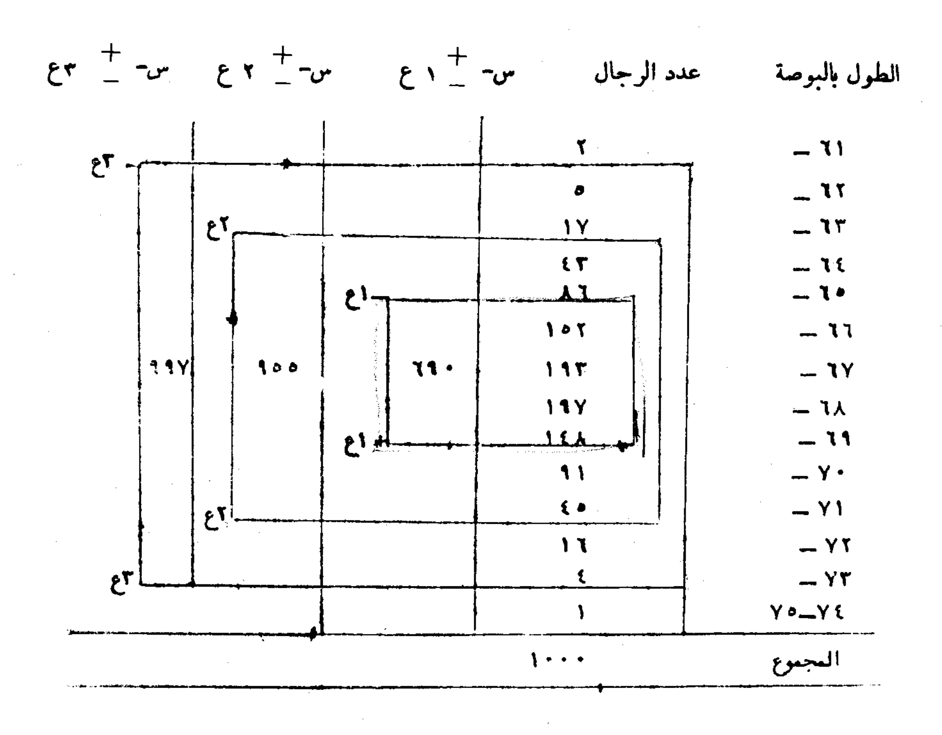
معنى التشتت في التوزيع التكراري :

نلاحظ أن التوزيعات التكراريه المعتدلة أو القريبة من الاعتدال (المتماثلة) لها خاصية تفيدنا كثيراً في الدراسات الاحصائية (العينات) .

ذلك أن المساحة المحدودة بالمنحنى التكراري المعتدل يمكن تقسيمها الى ثلاث قطاعات ، القطـــاع الأول يشمل ٦٨ر٢٧ ٪ من مجموع الوحدات (التكرارات) وقيمة كل وحدة منها لا تختلف عن الوسط الحسابي للتوزيع في الحـــد الأقصى الا بمقدار الانحراف المعيارى سواء بالزيادة أو بالنقص (س- — ع ، س- + ع) . أما القطاع الثاني فيشمل ٩٥ر٤٥ ٪ من مجموع الوحدات وقيمة كل وحدة منها لا تختلف عن الوسط الحسابي للتوزيع في الحد الأقصى الا بمقدار ضعف الانحراف المعياري سواء بالزيادة أو بالنقص (س- — ٢ع ، س- + ٢ع) . والقطاع الثاث ويشمل ٩٩ر٧٣ ٪ أي تقريباً مجموع الوحدات التي لا يمكن أن تختلف قيمة الوحدة منها عن الوسط الحسابي الا بمقدار ثلاثة أمثـــال الانحراف المعياري سواء بالزيادة أو بالنقص (س- — ٣ع ، س- + ٣ع) .

وبذلك نستطيع أن نقول أنه قلمـــا نجد وحدة واحدة في التوزيع التكراري المعتدل تختلف قيمتها عن الوسط الحسابي للتوزيع الا بمقدار ٣ع سواء بالزيادة أو بالنقص .

ويمكن توضيح الظاهرة السابقة بالجدول الآتي الذي يعطي مثلاً لتوزيع تكراري معتدل لأطوال مجموعة من الرجال بالبوصة .



الطول بالبوصة	عدد الرجال	س- ± ١ع	س- ± ٢ع	س- ± ٣ع
٦١ ـ	٢			
٦٢ ـ	٥			
٦٣ ـ	١٧			
٦٤ ـ	٤٣			
٦٥ ـ	٨٦			
٦٦ ـ	١٥٢			
٦٧ ـ	١٩٣	٦٩٠	٩٥٥	٩٩٧
٦٨ ـ	١٩٧			
٦٩ ـ	١٤٨			
٧٠ ـ	٩١			
٧١ ـ	٤٥			
٧٢ ـ	١٦			
٧٣ ـ	٤			
٧٤ـ٧٥	١			
المجموع	١٠٠٠			

ويمكن توضيح ذلك بيانياً كالآتي :

٪٩٩,٧٣ ٪٩٥,٤٥ ٪٦٨,٢٧

س-+٣ع س-س-—٣ع س—+٢ع س-س—+٢ع س-+١ع س-س-—١ع

وبمعنى آخر يمكن تقسيم مساحة المنحنى المعتدل (مجموع التكرارات في التوزيع المعتدل) الى قطاعات بدلالة الوسط الحسابي والانحراف المعياري لهذا التوزيع . فاذا كانت قيمة معينة في التوزيع تزيد عن الوسط الحسابي بمقدار ١ ع تكون المساحة المحصورة بين هذه القيمة وخط التماثل (الذي يمثل الوسط الحسابي للتوزيع) ٣٤ ٪ تقريباً من مجموع المساحة أي من مجموع التكررات . واذا كانت القيمة تنقص عن الوسط الحسابي بمقدار ١ ع تكون المساحة المحصورة بين هذه القيمة وخط التماثل ٣٤ ٪ من مجموع المساحة ، أي أن المساحة المحصورة بين القيمتين تكون ٦٨ ٪ من مجموع المساحة تقريباً .

واذا كانت القيمة تزيد عن الوسط الحسابي بمقدار ٢ ع فان المساحة المحصورة بين هذه القيمة وخط التماثل تكون ٤٧٫٥ ٪ تقريباً من مجموع المساحة . كذلك اذا كانث القيمة تقل عن الوسط الحسابي بمقدار ٢ ع تكون المساحة المحصورة بينها وبين خط التماثل ٤٧٫٥ ٪ أيضاً وبذلك يكون مجموع المساحة المحصورة بين القيمتين ٩٥ ٪ تقريباً من مجموع المساحة أي من مجموع التكرارات .

واذا كانت القيمة تزيد عن الوسط الحسابي بمقدار ٣ ع تكون المساحة المحصورة بينها وبين خط التماثل ٤٩٫٨٦ ٪ تقريباً من مجموع المساحة . كذلك اذا كانت القيمة تقل عن الوسط الحسابي بمقدار ٣ ع تكون المساحة المحصورة بينها وبين خط التماثل ٤٩٫٨٦ ٪ أيضاً . وبذلك يكون مجموع المساحة المحصورة بين القيمتين ٩٩٫٧٣ ٪ من مجموع المساحة أي من مجموع التكرارات .

وحتى نحدد كم يكون الفرق بين القيمة والوسط الحسابي من وحدات ع

نستخدم القانون $\frac{س - \bar{س}}{ع}$ (س رمز يدل على القيمة) ، تتوقف المساحة بين القيمة وخط التماثل على نتيجة هذا الكسر فاذا كانت ١ كانت المساحة ٣٤ ٪ وإذا كانت ٢ كانت ٤٧٫٥ ٪ وإذا كانت ٣ كانت المساحة ٤٩٫٨٦ ٪ من مجموع المساحة .

علاقات هامة بين مقاييس التشتت:

ونلاحظ انه بالنسبة للتوزيعات التكرارية القريبة من الاعتدال يكون الانحراف المتوسط $\frac{٤}{٥}$ الانحراف المعياري ، كما يكون نصف المدى الربيعي $\frac{٢}{٣}$ الانحراف المعياري . وتترتب هذه العلاقات على خاصية التوزيع المعتدل حيث يكون الانحراف المتوسط ٠٫٧٩٧٩ من الانحراف المعياري ، كما يكون نصف المدى الربيعي ٠٫٦٧٤٥ من الانحراف المعياري .

تمارين :

١ — احسب الانحراف المعياري للقيم : ١٢ ، ٦ ، ٧ ، ٣ ، ١٥ ، ١٠ ، ١٨ ، ٥ ، بثلاث طرق ع = ٤٫٨٧

كم يكون التباين لهذه القيم ؟

٢ – احسب الانحراف المعياري للتوزيع الآتي بطريقتين :

ف	ك
٦٠ – ٦٢	٥
٦٣ – ٦٥	١٨
٦٦ – ٦٨	٤٢
٧٢ – ٧٤	٨

ع = ٢٫٩٢

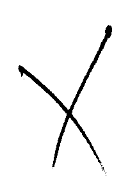
كم يكون التباين لهذا التوزيع ؟

٣ – احسب الانحراف المعياري للتوزيع الآتي :

معدلات الذكاء	٧٠	٧٤	٧٨	٨٢	٨٦	٩٠	٩٤	٩٨	١٠٢	١٠٦	١١٠	١١٤
عدد التلاميذ	٣٠	٩	١٦	٢٨	٤٥	٦٦	٨٥	٧٢	٥٤	٣٨	٢٧	١٨

احسب عدد التلاميذ الذين تتراوح معدلات ذكائهم بين س $\pm$ ١ ع .

» » » » » » » » س $\pm$ ٢ ع .



الفصل السابع

الالتـــواء

التواء أي توزيع يمثل انعدام التماثل فيه فجميع التوزيعات المتماثـــلة أو الطبيعية ينعدم فيها الالتواء نظراً لانتظامها أو تماثلها حول نقطة التركز فيها ، فوجود الالتواء دليل انعدام الانتظام في التوزيـع .

والالتواء أنواع فمن التوزيعات ما كان التواؤه معتدلاً ومنها ما كان التواؤه حاداً ومنها ما كان التواؤه موجباً ومنها كان التواؤه سالباً ، ولقد أشرنا الى جميع تلك الأنواع عند الكلام على أشكال المنحنيات التكرارية .

ويمكن الوقوف على طبيعة ودرجـــة التواء أي توزيع بمجرد النظر الى شكله البياني ولكن كثيراً ما نحتاج لتقدير درجة الالتواء بدقة فلا بد إذن من الاهتداء الى مقياس دقيق لهذه الظاهرة الهامة .

وسنقسم مقاييس الالتواء الى نوعين أساسيين ، الأول يصلح في جميـع التوزيعات التي يمثل الوسط الحسابي نقطة التركز فيها ، والثاني هو ما كان الوسيط أصلح وأنسب في تمثيلها .

مقاييس الالتواء

أولاً – حالات يمثل الوسط الحسابي فيها التوزيع

اذا كان التوزيع التكراري متماثلاً فان الوسط الحسابي والوسيط والمنوال تتساوى جميعها عند القيمة المقابلة لقمة المنحنى ، وكلما بعد التوزيع عن التماثل كلما اختلفت هذه المتوسطات الثلاث عن بعضها . وعلى ذلك يمكن أن نأخذ الفرق بين هذه المتوسطات كمقياس للالتواء . ولكن يكون من الأفضل أن نحصل على مقياس يمكن استخدامه في المقارنة ولهذا يقترح بيرسون أن نقسم الفرق بين الوسط والمنوال على الانحراف المعياري للتوزيع حيث أننا نعتبر الانحراف المعياري مقياس مدى ابتعاد القيم عن وسطها الحسابي .

وبذلك يكون مقياس بيرسون الالتواء = الوسط الحسابي – المنوال

$$\text{ومعامل بيرسون الالتواء} = \frac{\text{الوسط الحسابي – المنوال}}{\text{الانحراف المعياري}}$$

فاذا كان التوزيع متماثلاً فان الوسط الحسابي يتساوى مع المنوال وبذلك يكون مقياس الالتواء = صفر وكذلك معامل الالتواء . أما اذا كانت النتيجة أكثر من صفر دل ذلك على وجود التواء في التوزيع ودرجة هذا الالتواء تتوقف على قيمة المعامل واشارته الجبرية . فاذا كان معامل الالتواء صغيراً دل ذلك على ضعف الالتواء والعكس اذا كبر المعامل فان ذلك يدل على شدة الالتواء . واذا كان الوسط الحسابي أكبر من المنوال فان المعامل يكون موجباً وهذا دليل على التواء التوزيع نحو اليمين بمعنى ان القيم التي تزيد عن المنوال تكون تكراراتها أكبر من تكرارات القيم التي تقل عن المنوال . أما إذا كان الوسط الحسابي أقل من المنوال فان المعامل يكون

سالباً وهذا دليل على التواء التوزيع نحو اليسار بمعنى ان القيم التي تزيد عن المنوال تكون تكراراتها أقل من تكرارات القيم التي تقل عن المنوال .

وعيب معامل الالتواء السابق أنه يعتمد على المنوال وهو مقياس غير دقيق لا يجب أن نعتمد عليه في قياس درجة الالتواء . وخير الوسائل للاستغناء عنه التعبير عن الفرق بين الوسط الحسابي والمنوال بدلالة الوسيط بدلاً من المنوال . ويذكر القارىء ان هنالا علاقة تقريبية بين المتوسطات الثلاث في حالة التوزيعات المعتدلة الالتواء وهي :

الوسط الحسابي – المنوال = ٣ (الوسط الحسابي – الوسيط)

ولكن المهم أن لا نستخدم هذه العلاقة إلا في حالة التوزيعات القريبة من لتماثل أي المعتدلة الالتواء . ويكون معامل الالتواء في هذه الحالة :

$$\text{معامل الالتواء} = \frac{\text{٣ (الوسط الحسابي – الوسيط)}}{\text{الانحراف المعياري}}$$

على أنه يجب أن نلاحظ أن معامل الالتواء المحسوب على الأساس الأول يختلف بعض الشيء عن المعامل المحسوب على الأساس الثاني وذلك لأن العلاقة بين المقياسين علاقة تقريبية

ثانياً – حالات يمثل الوسيط فيها التوزيع

اذا كان التوزيع مفتوحاً فان بولي Bowley يقترح مقياساً آخر للالتواء وهو في نفس الوقت اسهل في الحساب من مقياس بيرسون Pearson . ويقوم مقياس بولي على أساس العلاقة بين الربيع الأعلى والأدنى والوسيط، ذلك لأنه في حالة التوزيع التكراري المتماثل تتساوى المسافة بين كل من الربيعين

والوسيط ، ولا يتحقق ذلك إذا كان التوزيع ملتوياً . وعلى ذلك يمكننا قياس الالتواء بطرح الفرق بين الربيع الأعلى والوسيط من الفرق بين الوسيط والربيع الأدنى . ذلك لأنه اذا كان التوزيع متماثلاً فان الفرق يكون صفراً أما اذا كان الفرق أكبر من صفر فان هذا دليل على وجود التواء في التوزيع تتوقف درجته ونوعه على قيمة الفرق واشارته الجبرية ، فاذا كان الفرق كبير كان ذلك دليلاً على وجود التواء شديد والعكس اذا كان الفرق بسيطاً ، وإذا كان الفرق موجباً كان دليلاً على التواء التوزيع نحو اليمين والعكس إذا كان الفرق سالباً كان ذلك دليلاً علي الالتواء نحو اليسار .

ولما كان هذا الفرق رقماً مميزا فانه لا يصلح للمقارنة ولهذا يقترح بولى قسمة هـذا الفرق على مجموع المسافة بين كل من الربيعين والوسيط أي على الانحراف الربيعي :

وعلى ذلك يكون مقياس بولى للالتواء = (الربيع الأعلى – الوسيط) – (الوسيط – الربيع الأدنى)

ويكون معامل بولى للالتواء = (الربيع الأعلى – ٢ الوسيط – الربيع الأدنى) ÷ (الربيع الأعلى – الربيع الأدنى)

وميزة هذا المقياس هو سهولة حساب الوسيط والربيعين بالنسبة لصعوبات حساب الانحراف المعياري والوسط الحسابي والمنوال .

على انـه يجب أن نلاحظ أن النتيجة التي نحصل عليها بتطبيق هذا المعامل تختلف عن تلك التي نحصل عليها اذا استخدمنا معامل بيرسون ، وذلك لأن كل من المقياسين يقوم على أساس مختلف ، فلا يجب استخدام المقياسين لمقارنة التواء توزيعين تكرارين وانما يجب استخدام نفس المقياس اذا أردنا المقارنة .

العزوم ومقاييس الالتواء :

عزم أي قوة هو مقدار العمل الذي تحدثه ، ويتوقف هذا العمل على القوة نفسها والمسافة بين هذه القوة والنقطة التي عندها تحدث اثرها ، فقوة مقدارها ٨ كيلو على بعد ١ قدم من نقطة الأصل (الصفر) تعادل في مفعولها قوة مقدارها ٢ كيلو على بعد ٤ قدم من نقطة الاصل ، أي أن التوازن يتحقق عند تساوي الناتج الموجب مع النانج السالب .

وبالنسبة للتوزيع التكراري تكون تكرارات التوزيع هي القوى المؤثرة عليه وعزم أي تكرار يقاس بحاصل ضرب التكرار في انحرافه عن نقطة الأصل في التوزيع التي يعبر عنها الوسط الحسابي . والعزم الأول هو حاصل جمع عزوم تكرارات التوزيع مقسوماً على مجموع التكرارات في التوزيع أي $\frac{\text{مجـ ح ك}}{\text{مجـ ك}}$ ويعبر عنه في العزوم بالرمز مـ$_{1}$ وحيث أن مجموع الانحرافات عن الوسط الحسابي = صفراً ، لذلك يكون العزم الأول في التوزيع التكراري مساويا صفراً دائماً . أما اذا استبدلنا العزم الاول عن الوسط الحسابي بالعزم الاول عن القيمة صفر فمن الواضح أن الناتج يكون هو الوسط الحسابي نفسه .

أي ان العزم الأول (حول الصفر) = $\frac{\text{مجـ (س – صفر)ك}}{\text{مجـ ك}}$

= الوسط الحسابي .

لاحظنا أننا في هـذه الحالة كأننا حسبنا الوسط الحسابي بدلالة وسط وسط فرضي صفر وبذلك يكون الوسط الحسابي مساوياً لمجموع الانحرافات عن الصفر مقسوماً على مجموع التكرارات (ارجع الى قانون الوسط الحسابي بدلالة الوسط الفرضي) .

والعزم الثاني للتوزيع التكراري وهو متوسط مربعات انحرافات القيم في التوريع (المراكز) عن الوسط الحسابي أي :

$$\text{م}_{٢} = \frac{\text{مج ح}^{٢}\text{ ك}}{\text{مج ك}}$$

ومن الواضح أن العزم الثاني يعطينا التباين أي مربع الانحراف المعياري .

والعزم الثالث التوزيع التكراري وهو متوسط مكعبات انحرافات القيم في التوزيع (المراكز) عن الوسط الحسابي أي :

$$\text{م}_{٢} = \frac{\text{مج ح}^{٣}\text{ ك}}{\text{مج ك}}$$

ويكون هذا العزم مساوياً صفر في التوزيع الطبيعي فقط ، أما في التوزيع الملتو يكون العزم قيمة موجبة أو سالبة تبعاً لنوع الالتواء ، ذلك لأن تكعيب الانحرافات لا يخلصنا من الاشارة الجبرية فتظل انحرافات القيم الأقل من المتوسط تحمل اشارة مضادة لانحرافات القيم الاكبر عن المتوسط وبذلك عند الجمع يكون الناتج اما موجباً أو سالبا ، فإذا كانت الانحرافات الكبيرة سالبة يكون المجموع الجبري لمجموع مكعبات الانحرافات سالباً ولو كانت الانحرافات الكبيرة موجبة فان المجموع الجبري لمكعبات الانحرافات يكون موجباً . وبذلك يمكن استخدام خاصية العزم الثالث في قياس درجة التواء التوزيع . وحيث اننا عند قياس الالتواء يكون هدفنا غالباً هو مقارنة التواء توزيعات مختلفة فلا يمكن الاعتماد على العزم الثالث في المقارنة ، ولذلك لكي نحصل على مقياس نسبي للالتواء نقسم العزم الثالث على مكعب الانحراف المعياري حيث أن وحدات القياس في البسيط تكون مكعبة ولذلك يجب القسمة على قيمة مرفوعة الى نفس القوة في المقام .

وبذلك يكون معامل الالتواء $B_1 = \frac{م_3}{(\sqrt{م_2})^3}$ أي $\frac{م_3}{ع^3}$

وقد اقترح بيرسون لأسباب رياضية استخدام مربع المعامل السابقة كمقياس للالتواء : —

ولذلك يكون $B_1 = \frac{(م_3)^2}{(م_2)^3}$

وحيث ان الانحرافات عن الوسط الحسابي يمكن ان تكون أرقام كبيرة وتكعيبها يؤدي الى ضياع وقت كبير ، لذلك يحسب العزم الثالث على أساس وسط فرضي ثم يصحح حتى نحوله الى العزم الثالث على أساس الوسط الحسابي باستخدام العلامة الآتية : —

$$م_3 = \frac{مج\ ح^3\ ك}{مج\ ك} - 3\ \frac{مج\ ح\ ك}{مج\ ك} \times \frac{مج\ ح^2\ ك}{مج\ ك} + 2\left(\frac{مج\ ح\ ك}{مج\ ك}\right)^3$$

مثال ٢٤ :

ف	ك	ح	ح ك	$ح^2$ ك	$ح^3$ ك
١٢	١	-٣	-٣	٩	-٢٧
١٤	٤	-٢	-٨	١٦	-٣٢
١٦	٦	-١	-٦	٦	-٦
١٨	١٠	صفر	صفر	صفر	صفر
٢٠	٧	١٠	+٧	٧	+٧
٢٢	٢	+٢	+٢		
	٣٠		-٨	٤٢	-٥٠

$$\text{مـ}_٣ = \frac{\text{مجـ ح}^٣\text{ك}}{\text{مجـ ك}} - ٣ \frac{\text{مجـ ح ك}}{\text{مجـ ك}} \times \frac{\text{مجـ ح}^٢\text{ك}}{\text{مجـ ك}} + ٢ \left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}} \right)^٣$$

$$= \frac{-٥٠}{٣٠} - ٣ \times \frac{-٨}{٣٠} \times \frac{٤٢}{٣٠} + ٢ \times \left(\frac{٨٠}{٣٠} \right)^٣$$

$$= -١ر٦٦٦٧ + ١ر١٢ - ر٠٣٧٩$$

$$= -٠ر٥٨٣٦$$

وإذا أردنا معرفة العزم الثالث بوحدات طول الفئة تضرب النتيجةالسابقة في مكعب طول الفئة على انه عند حساب الالتواء لا داعي لذلك حيث يمكن ان نقسم العزم الثالث على مكعب الانحراف المعياري (بـــدون الضرب في طول الفئة) ثم تربيع الناتج نحصل على معامل بيرسون للالتواء .

وبذلك يكون ع بدون التحويل الى وحدات القياس الخاصة بطول الفئة .

$$= \sqrt{\frac{٤٢}{٣٠} - \left(\frac{٨}{٣٠} \right)^٢}$$

$$\therefore \text{ع}^٣ = ١ر٥٢٩٥$$

$$\therefore \text{معامل الالتواء} = \left(\frac{-٠ر٥٨٣٦}{١ر٥٢٩٥} \right) = -ر٣$$

وبالمثل يمكن حساب العزم الرابع حول الوسط الحسابي وهو عبارة عن متوسط مجموع انحرافات القيم عن الوسط الحسابي مرفوعة الى القوة الرابعة ويمكن التعبير عنه رمزيا

$$م_4 = \frac{مج ح^4 ك}{مج ك}$$

وعلى نفس النمط يمكن قياس العزم الخامس والسادس . . الخ ، إلا أن العزوم الأربعة الاولى هي الشائعة الاستعمال في الاحصاء ، حيث أن العزم الثاني يدلنا على التباين والعزم الثالث يدلنا على الالتواء والعزم الرابع يدلنا على التفرطح . والتفرطح يعني ان قمة المنحنى الذي يمثل التوزيع مفرطحة وليست مدببة ، ولا شك انه كلما كان تشتت الوحدات حول متوسطها صغيراً كلما قل التفرطح والعكس كلما كان التشتت كبيرا كلما زاد التفرطح ، إلا انه لا توجد أي صلة بين التشتت والتفرطح فكل منهما يصف خاصية مختلفة تماماً في التوزيـع التكراري فالتفرطح يُعنى به تفرطح قمة المنحنى وليس انفراج المنحنى من أسفل أي اتساع المدى بين طرفيه على المحور الافقي وهو ما يدل على تشتت التوزيع . والتوزيـع الطبيعي بالرغم من انه يكون عادة متوسط التفرطح الا أن درجة تفرطحه تختلف من منحنى إلى آخر .

تمـــارين :

١ — احسب معامل الالتواء للتوزيع الآتي بالطرق المختلفة :

الفئات	التـكرار
٩٫٣ — ٩٫٧	٢
٩٫٨ — ١٠٫٢	٥
١٠٫٣ — ١٠٫٧	١٢
١٠٫٨ — ١١٫٢	١٧
١١٫٣ — ١١٫٧	١٤
١١٫٨ — ١٢٫٢	٦
١٢٫٣ — ١٢٫٧	٣
١٢٫٨ — ١٣٫٢	١

٢ – إذا كان العزم الثاني لتوزيعين تكراريين هو ٩ ، ١٦ على التوالِ ، والعزم الثالث لنفس التوزيعين هو – ١٫٨ ، – ١٢٫٨ ، أي من التوزيعين أكثر التواء وما هو اتجاه التواء التوزيعين . التوزيـع الاول أكثر التواء .



الفصل الثامن

توفيق المنحنى الطبيعي (المعتدل)

أشرت سابقاً إلى بعض خواص التوزيع الطبيعي (المعتدل) ، ونظراً لأهمية المنحنى في الدراسة الخاصة بالعينات يكون من المفيد ان ندرس خواصه دراسة تفصيلية .

لقد تبين لنــــا من الفصل السادس كيف يمكن تقسيم مساحة المنحنى الطبيعي الى قطاعات محددة بدلالة الوسط الحسابي والانحراف المعياري للتوزيع الذي يمثله هذا المنحنى . وبمعنى آخر يمكننا أن نعين نسبة مساحة المنحنى الطبيعي المحصورة بين قيمتين من قيم الظاهرة التي يمثلها المنحنى إلى المساحة الكلية اذا عرفنا انحراف كل من القيمتين عن الوسط الحسابي مقدراً بوحدات الانحراف المعياري . فاذا كان انحراف كل من القيمتين عن الوسط الحسابي للتوزيع يساوي + ١ ع أو - ١ ع ، فان المساحة المحصورة بين هاتين القيمتين تمثل ٦٨ر٢٧ ٪ من المساحة الكلية للمنحنى أي من مجموع الوحدات التي يمثلها التوزيع . أما إذا كان انحراف كل من القيتين عن الوسط الحسابي للتوزيع يساوي + ٢ ع أو - ٢ ع ؛ فإن المساحة المحصورة بـــين هاتين القيمتين تمثل ٩٥ر٤٥ ٪ من مجموع المساحة ، وإذا كان انحراف كل من القيمتين عن الوسط الحسابي للتوزيع يساوي + ٣ ع أو - ٣ ع فان المساحة المحصورة بــــين القيمتين تمثل ٧٣ ر ٩٩ ٪ من مجمــوع المساحة الكلية للمنحنى

على أن تقسيم مساحة المنحنى الطبيعي بدلالة الوسط الحسابي والانحراف المعياري لا يقتصر فقط على هذه القطاعات الثلاث الأساسية ، إذ أمكن التوسع في تطبيق هذه الخاصة باعداد جدول مساحات المنحنى الطبيعي ومنه يمكن تحديد نسبة المساحة المحصورة بين أي قيمة وبين خط التماثل من مجموع المساحة الكلية إذا استطعنا تحديد انحراف هذه القيمة عن الوسط الحسابي للتوزيع بدلالة وحدات الانحراف المعياري ، ويكون ذلك بقسمة الفرق بين القيمة وبين الوسط الحسابي على الانحراف المعياري وبذلك نستطيع أن نعرف مقدار هذا الفرق بدلالة وحدات الانحراف المعياري ، أي هل يساوي الفرق ± ١ ع أو ± ١ر٢ ع أو ± ١ر٥ ع أو ± ١ر٧ ع أو ± ٢ر٢ ع، فإذا حددنا ذلك نستطيع ان نحدد المساحة المطلوبة كنسبة من المساحة الكلية .

وحيث ان كثيراً من الدراسات الإحصائية تقوم على أساس المعاينة وحيث ان نتائج العينة لا يمكن ان نطابق تماماً صفات المجتمع الإحصائي الذي سحبت منه العينة ، لذلك قد يحدث أحياناً انه بعد تبويب البيانات الخاصة بالعينة في توزيع تكراري يتراءى لنا أن هناك ما يدعو الى الاعتقاد بأن هذه البيانات موزعة توزيعاً طبيعياً خاصة إذا أعدنا التجربة على عينات أخرى ووجدنا أنه في كل مرة توزع قيم الظاهرة توزيعاً يقرب من الاعتدال وأنه إذا كبرنا حجم العينة قرب التوزيع من الوضع المعتدل . وبذلك يمكن في هذه الحالة أن نستنتج أنه لو كان من الممكن عملياً جمع البيانات من جميع وحدات المجتمع الإحصائي موضوع الدرس لوجدنا انها موزعة فعلاً توزيعاً معتدلاً . ولما كانت دراسة العينات هي الوسيلة لدراسة المجتمعات التي سحبت منها فيمكن أن نكتفي بعينة واحدة على أن نمهد المنحنى الطبيعي الذي يمثل التوزيع أي نجعله يتخذ الوضع المثالي الذي يكون عنده معتدلاً تماماً . ولزيادة الإيضاح نقول ان دراسة عينة من مجتمع ما لا يمكن أن يظهر فيها

نفس التوزيع النسبي لمختلف أنواع الوحدات التي تكوّن هذا المجتمع ، فبعض الأنواع قد تظهر في العينة بوزن يزيد عن وزنها في المجتمع ، ولهذا إذا كان توزيع وحدات المجتمع توزيعاً طبيعياً فإن ذلك لن يظهر تماماً في العينة وبمعنى آخر إذا رسمنا المنحنى التكراري الذي يمثل وحدات العينة فلن يكون شكله معتدلاً مائة في المائة وإنما ستظهر فيه بعض التعرجات الناتجة عن ظهور بعض الوحدات في العينة بنسبة أكبر من نسبتها في المجتمع وظهور البعض الآخر بنسبة أقل من نسبتها في المجتمع . إلا أن هـــذه التعرجات في المنحنى قد تكون من الصغر بحيث يمكن أن نحكم بأن المنحنى لو ظهرت وحدات العينة بنفس توزيعها النسبي في المجتمع يمكن أن يكون منحنى تام الاعتدال ولهـذا نقول « كيف يمكن أن يكون شكل المنحنى الذي يمثل العينة لو ظهرت الوحدات فيها بنفس توزيعها النسبي في المجتمع الذي تمثله ؟ وهذا القول هو الذي يدعونا الى توفيق المنحنى الذي يمثـــل وحدات العينة أي القضاء على التعرجات الشاذة التي تظهر فيه .

وعند توفيق المنحنى الذي يمثل وحدات العينة لا بد ان يتفق المنحنى بعد التوفيق مع المنحنى قبل التوفيق في المجموع الكلي للوحدات وفي الوسط الحسابي لقيمتهـــا وكذلك في انحراف هذه القيم عن الوسط الحسابي أي في انحرافها المعياري . فنحن وإن كنا نغيّر من شكل المنحنى أي من توزيـع الوحدات في الفئات المختلفة فلا يجب أن يتميز بذلك مجموع التكرارات أو الوسط الحسابي أو الانحراف المعياري للتوزيع الاصلي الذي نقوم بتوفيقه لأن تغيير هذه المقاييس يعني تغيير التوزيع تغييراً جذريا وهذا ليس المقصود من عملية التوفيق ولهذا تؤخذ هذه المقاييس كدعائم يرتكز عليها عملنا الحسابي الذي تتضمنه عملية التوفيق .

والمقصود بعملية التوفيـــق هو حساب التكرارات النظرية التي تقابل التكرارات الواقعية أي المشاهدة نتيجة جمع المعلومات الاحصائية عن الظاهرة

موضوع الدرس . ويمكن إجراء عملية التوفيق بطريقتين ، امـــا عن طريق خاصية التوزيع الطبيعي التي سبق الاشارة اليها وهي امكانية تقسيم مساحته الى قطاعات معلوم نسبتها الى المساحة الكلية وفي هذه الحالة نستخدم جدول مساحات المنحنى الطبيعي ، واما عن طريق خاصية اخرى وهي إمكانية تحديد الاحداثيات الرأسية للمنحنى المعتدل .

توفيق المنجنى بطريقة المساحات :

اشرت في الفقرة السابقة الى الملاحظتين لآتيتين : —

١ — ان عملية التوفيق لا يجب أن تغير من مجموع التكرارات أو الوسط الحسابي أو الانحراف المعياري للتوزيع الأصلي ولهذا نحسب هذه المقاييس اولاً وتؤخذ كدعائم يرتكز عليها عملنا الحسابي .

٢ — إن تطبيق هذه الطريقة يحتاج الىاستحدام جدول مساحات المنحنى الطبيعي ، وهذا الجدول يعطي نسبة المساحة المحصورة بين الاحداثي الرأسي المتوسط للتوزيع (أي الذي يقابل قمة المنحنى والوسط الحسابي على المحور الأفقي) وبين الأحداثي الرأسي المقام على أي نقطة تمثل أي قيمة على يمين أو على يسار الوسط الحسابي للتوزيع . والجدول يعطي هذه النسبة إذا علمنا نسبة الفرق بين القيمة والوسط الحسابي للتوزيع الى الانحراف المعياري له أي $\frac{س - \bar{س}}{ع}$ وبمعنى آخر إذا علمنا كم يساوي الفرق بين القيمة والوسط الحسابي من وحدات الانحراف المعياري للتوزيع . ويجب أن نلاحظ أن الجدول يعطي هذه النسبة على أساس ان المجموع الكلي لمساحة المنحنى واحد صحيح ؛ فاذا أردنا هذه النسبة في كل شكل مئوي فما علينا إلا أن نضرب النسبةالتي تظهر في الجدول في ١٠٠ . كذلك يجب أن نلاحظ أن الجدول يعطي النسبة على أساس وجود القيمة على يمين أو على يسار محور التماثل .

مثـــال ٢٥ :

وفيما يلي اورد المثال التالي لايضاح طريقة العمل :

حسبت الأجور اليومية لألف عامل وبوبت في الجدول الآتي :

فئات الأجر بالليرة	عدد العمال
٤ر٥ –	٦
٥ر٥ –	٧
٦ر٥ –	٣٥
٧ر٥ –	٤٨
٨ر٥ –	٦٥
٩ر٥ –	٩٠
١٠ر٥ –	١٣١
١١ر٥ –	١٧٣
١٢ر٥ –	١٥٥
١٣ر٥ –	١١٧
١٤ر٥ –	٧٥
١٥ر٥ –	٥٢
١٦ر٥ –	٢١
١٧ر٥ –	٩
١٨ر٥ – ١٩ر٥	٦

واقترح احد الاحصائيين أن هذه البيانات (على أساس عينة من ١٠٠٠

عامل) توحي بأن التوزيع الذي يمثل هذه الظاهرة هو توزيع طبيعي ، ولذلك نبـــدأ بتوفيق المنحنى الطبيعي أي بحساب تكرار كل فئة لو كان التوزيع طبيعياً تماماً .

خطوات العمل :

١ ـ حساب الوسط الحسابي والانحراف المعياري للتوزيـع الواقعي .

٢ ـ حساب الفرق بين كل قيمة (س) تمثل الحدود الدنيا للفئات والوسط الحسابي للتوزيع .

٣ ـ حساب نسبة هذه الفروق الى الانحراف المعياري أي حساب الدرجـــة المعيارية $\frac{س - \bar{س}}{ع}$.

٤ ـ استخراج نسبة المساحة المحصورة بين كل قيمة ومحور التماثل في التوزيع من جدول المساحات للتوزيع الطبيعي .

٥ ـ طرح كل نسبة من التي تليها لتحديد نسبة مساحة كل فئة الى المجموع الكلي للمساحة .

٦ ـ لما كانت المساحة الكلية للتوزيع يمثلها مجموع التكرارات ، لذلك تضرب نسبة مساحة كل فئة في مجـــوع التكرارات نحصل على التكرار النظري لكل فئة .

ويمكن اجراء الخطوات السابقة في الجدول الآتي :ــ

الفئات	ك	م	ح	ح ك	ح² ك	س ــ س̄	(س ــ س̄) / ع	المساحة بين س ومحور التماثل %	مساحة الفئة %	تكرار الفئة
٤ر٥ـ	٦	٥	ـ٧	ـ٤٢	٢٩٤	٧ر٥٠٦	٢ر٨٦	٤٩ر٧٩	٤٥ر	٦
٥ر٥ـ	١٧	٦	ـ٦	ـ١٠٢	٦١٢	٦ر٥٠٦	٢ر٤٨	٤٩ر٣٤	١ر١٣	١٢
٦ر٥ـ	٣٥	٧	ـ٥	ـ١٧٥	٨٧٥	٥ر٥٠٦	٢ر١	٤٨ر٢١	٢ر٥٧	٢٦
٧ر٥ـ	٤٨	٨	ـ٤	ـ١٩٢	٧٦٨	٤ر٥٠٦	١ر٧١	٤٥ر٦٤	٤ر٨١	٤٨
٨ر٥ـ	٦٥	٩	ـ٣	ـ١٩٥	٥٨٥	٣ر٥٠٦	١ر٣٣	٤٠ر٨٣	٧ر٩٤	٧٩
٩ر٥ـ	٩٠	٤	ـ٢	ـ١٨٠	٣٦٠	٢ر٥٠٦	٩٥ر	٣٢ر٨٩	١١ر٣٢	١١٣
١٠ر٥ـ	١٣١	١١	ـ١	ـ١٣١	١٣١	١ر٥٠٦	٥٧ر	٢١ر٥٧	١٤ر٠٤	١٤٠ر٥
١١ر٥ـ	١٧٣	١٢	صفر	صفر	صفر	٥٠٦ر	١٩ر	ـ٧ر٥٣	١٥ر٠٦	١٥١
١٢ر٥ـ	٤٥٥	١٣	١	١٥٥	١٥٥	٤٩٤ر	١٩ر	+ ٧ر٥٣	١٤ر٠٤	١٤٠ر٥
١٣ر٥ـ	١١٧	١٤	٢	٢٣٤	٤٦٨	١ر٤٩٤	٥٧ر	٢١ر٥٧	١١ر٣٢	١١٣
١٤ر٥ـ	٧٥	١٠	٣	٢٢٥	٦٧٥	٢ر٤٩٤	٩٥ر	٣٢ر٨٩	٧ر٩٤	٧٩
١٥ر٥ـ	٥٢	١٦	٤	٢٠٨	٨٣٢	٣ر٤٩٤	١ر٣٣	٤٠ر٨٣	٤ر٨١	٤٨
١٦ر٥ـ	٢١	١٧	٥	١٠٥	٥٢٥	٤ر٤٩٤	١ر٧١	٤٥ر٦٤	٢ر٥٣	٢٦
١٧ر٥ـ	٩	١٨	٦	٥٤	٣٢٤	٥ر٤٩٤	١ر٠٩	٤٨ر١٧	١ر١٥	١٢
١٨ر٥ـ١٩	٦	١٩	٧	٤٢	٢٩٤	٦ر٤٩٤	٢ر٤٧	٤٩ر٣٢	٤٦ر	٦
١٩ر٥						٧ر٤٩٤	٢ر٨٥	٤٩ر٧٨		
	١٠٠٠			+ ٦	٦٨٩٨					

$$س̄ = ١٢ + \frac{٦}{١٠٠٠} ١ = ١٢ر٠٠٦ \text{ ليرة}$$

$$ع = ١ \sqrt{\frac{٦٨٩٨}{١٠٠٠} - \frac{٢٦}{١٠٠٠}}$$

$$= ١ \sqrt{٦ر٨٩٨ - ٠ر٠٠٠٠٣٦}$$

$$= ١ \sqrt{٦ر٨٩٧٩٦٤}$$

$$= ٢ر٦٢٦ \text{ ليرة .}$$

حصلنا في العمود الأخير من الجدول السابق على التكرارات النظرية أي للتكرارات التي يمكن توقعها لو كان التوزيع متماثلا تمام التماثل . وبالقاء نظرة على هذه التكرارات يتبين أن هناك فروق بينها وبين التكرارات الواقعية وهو أمر كنا نتوقعه . والمشكلة هي - هل هذه الفروق من الكبر بحيث يتعين علينا استبعاد الافتراض الذي بدأنا به وهو اعتدال المجتمع الذي يمثله هذا التوزيع أم انها من الصغر بحيث يمكن أن نعزوها إلى المصادفة الناتجة عن المعاينة العشوائية . والبث في هذه المشكلة يكون باجراء اختبار كاي تربيع الذي سنأتي إلى مناقشته في الفصل التالي :

توفيق المنحنى بطريقة الاحداثيات الرأسية :

استطاع الرياضيون أن يحددوا المعادلة التي تمثل المنحنى المعتدل وهي كالآتي :

$$ص = \frac{1}{ع\sqrt{2\,ط}} \, هـ^{\frac{-س^2}{2\,ع^2}}$$

حيث ص هي الاحداثي الرأسي المناظر لقيمة س .

، س هي الاحداثي الافقي مقيساً من نقطة الأصل التي تقع عند الوسط الحسابي للتوزيع وبعبارة اخرى فان قيم س تمثل الانحرافات عن الوسط الحسابي للتوزيع :

، ع هي الانحراف المعياري للتوزيع

، ط هى النسبة التقريبية (٣٫١٤١٦)

، هـ هي أساس لوغاريتمات نابيير (٢٫٧١٨٢٨) .

وإذا أردنا تحديد الاحداثي الرأسي لتوزيـع طبيعي معين عند أي قيمة على المحور الافقي يجب ضرب المعادلة السابقة في عدد وحدات التوزيـع ، وبمعنى آخر يستبدل العدد بوحدات التوزيـع . وبذلك يكون الاحداثي الرأسي المتوسط للتوزيـع أي الاحداثي المقام على الوسط الحسابي للتوزيـع =

$$\frac{١}{ع\sqrt{٢\,ط}}$$

وإذا كان مجموع التكرارات = مجـ ك يكون الاحداثي

الرأسي المتوسط = $$\frac{مجـ\ ك}{ع\sqrt{٢\,ط}}$$ حيث أن $هـ^{\frac{-س^{٢}}{٢\,ع^{٢}}}$ ، $هـ^{\frac{-س^{٢}}{٢\,ع^{٢}}}$

سـوف تساوي صفر لأن القيمة $-س^{٢}$ = صفر إذ أن انحرافها عن الوسط الحسابي ، يساوي صفر .

مثـــال ٣٦ :

منحنى معتدل يمثل ١٠٠٠ وحدة ، وانحرافه المعياري = ١٠ كم يكون التكرار الذي يدل عليه محور التماثل فيه :

$$ص = \frac{١٠٠٠}{١٠\sqrt{٢ \times ٣ر١٤١٦}} = ٣٩ر٨٩٤$$

وبذلك يمكن تحديد أي احداثي رأسي للمنحنى بضرب محور التماثل فيه في القيمة :

فالاحداثي الرأسي فوق النقطـــة التي تبعد عـــن الوسط

$- \frac{٤٠٠}{٢٠٠}$

الحسابي بمقدار ٢٠ = ٣٩٫٨٩٤ × ٢٫٧١٨٢٨

$= ٣٩٫٨٩٤ \times \frac{١}{٢٫٧١٨٢٨^{٢}} = ٥٫٣٩٩$.

ولتسهيل العمل عند تحديد الاحداثيات الرأسية التي تكون المنحنى المعتدل يمكن الرجوع إلى جدول احداثيات المنحنى الطبيعي . ويعطى هذا الجدول ارتفاع الاحداثي الرأسي عند أي نقطة تبعد عن الوسط الحسابي للتوزيع كنسبة مئوية من الاحداثي الرأسي المتوسط (خط التماثل) . والابتعاد عن الوسط الحسابي يكون بدلالة الانحراف المعياري للتوزيع أي $\frac{س - \bar{س}}{ع}$.

لذلك حتى يمكن استعمال هذا الجدول يجب تحديد الابتعاد عن الوسط الحسابي للتوزيع بدلالة الانحراف المعياري أي تحديد قيمة $\frac{س - \bar{س}}{ع}$ ،

والقيمة المقابلة في الجدول (أي القيمة المقابلة للدرجة المعيارية $\frac{س - \bar{س}}{ع}$) تضرب في الاحداثي الوأسي المتوسط :

ويحدد هذا الاحداثي باستخدام المعادلة $\frac{ح ك}{ع \sqrt{٢ \times ٢٫٢٤١٦}}$

أي : $\frac{ح ك}{ع \times ٢٫٥٠٦٦٢٨}$

ويمكن وضع هذه المعادلة في الصيغه :

$$\text{٠٫٣٩٨٩٤} \times \frac{\text{مج ك} \times \text{طول الفئة}}{\text{ع (بالوحدات الأصلية للتوزيع)}}$$

مثـــال ٢٧ :

منحنى طبيعي يمثل ٦٠٠ وحدة وانحرافه المعياري = ٠٫٠٠٨٥ ووسطه الحسابي = ٠٫٠٢٠٢ ، كم يكون المحور الرأسي للقيمة ٠٫٠١٨٢ إذا كان طول الفئة في هذا التوزيع = ٠٫٠٠٠٤ ؟

$$\text{س} - \bar{\text{س}} = \text{٠٫٠١٨٢} - \text{٠٫٠٢٠٢} = \text{٠٫٠٠٢}$$

$$\frac{\text{س} - \bar{\text{س}}}{\text{ع}} = \frac{\text{٠٫٠٠٢}}{\text{٠٫٠٠٨٥}} = \text{٢٫٣٥}$$

من جدول الاحداثيات الرأسية للمنحنى الطبيعي يكون الاحداثي الرأسي المقابل لهذه القيمة يشكل ٦٫٢١ ٪ من الاحداثي المتوسط .

والخطوة التالية هي حساب الاحداثي الرأسي المتوسط باستخـدام المعـــادلة :

$$\text{٠٫٣٩٨٩٤} \times \frac{\text{٦٠٠} \times \text{٠٫٠٠٠٤}}{\text{٠٫٠٠٨٥}} = \text{١١٢٫٦}$$

أو باستخـــدام المعادلة :

$$\frac{\text{٦٠٠}}{\text{٢٫١٢٥} \sqrt{\text{٢} \times \text{٣٫١٤١٦}}} = \text{١١٢٫٦}$$

ونلاحظ ان القيمة ٢٫١٢٥ هي الانحراف المعياري بدلالة طول الفئة في التوزيع أي بقسمة الانحراف المعياري بالوحـــدات الاصلية للتوزيع على طول الفئة .

وبذلك يكون الاحداثي الرأسي للقيمة ٠٫٠٠٨٤ =

$$\frac{٣٢١ر٦}{١٠٠} \times ١١٢ر٦ = ١٢ر٧ .$$

مثال ٢٨ :

ويمكن توضيح عمليه توفيق المنحنى بطريقــة الاحداثيات الرأسية بالجدول الآتي :

الفئات	ك	م	م – س	(م – س) / ع	نسبة الاحداثي الرأسي إلى الاحداثي المتوسط	احداثي المركز
٤ر٥ —	٦	٥	٧ر٠٠٦	٢ر٦٦	ر٠٢٩٠٨	٦
٥ر٥ —	١٧	٦	٦ر٠٠٦	٢ر٣٣	ر٠٦٦٢٤	١٢
٦ر٥ —	٣٥	٧	٥ر٠٠٦	١ر٩١	ر١٦١٣٧	٢٥
٧ر٥ —	٤٨	٨	٤ر٠٠٦	١ر٥٢	ر٣١٥٠٠	٤٨
٨ر٥ —	٦٥	٩	٣ر٠٠٦	١ر١٤	ر٦٢٢١٤	٧٩
٩ر٥ —	٩٠	١٠	٢ر٠٠٦	ر٧٦	ر٧٤٩١٦	١١٣
١٠ر٥ —	١٣١	١١	١ر٠٠٦	ر٣٨	ر٩٣٠٢٤	١٤١
١١ر٥ —	١٧٣	١٢	٠ر٠٠٦	صفر	١ر٠٠٠٠	١٥٢
١٢ر٥ —	١٥٥	١٣	٠ر٩٩٤	ر٣٨	ر٩٣٠٢٤	١٤١
١٣ر٥ —	١٦٧	١٤	١ر٩٩٤	ر٧٦	ر٧٤٩١٦	١١٣
١٤ر٥ —	٧٥	١٥	٢ر٩٩٤	١ر١٤	ر٥٢٢١٤	٧٩
١٥ر٥ —	٥٢	١٦	٣ر٩٩٤	١ر٥٢	ر٣١٥٠٠	٤٨
١٦ر٥ —	٢١	١٧	٤ر٩٩٤	١ر٩١	ر١٦١٣٧	٢٥
١٧ر٥ —	٩	١٨	٥ر٩٩٤	٢ر٣٣	ر٠٦٦٢٣	١٢
١٨ر٥ — ١٩ر٥	٦	١٩	٦ر٩٩٤	٢ر٦٦	ر٠٢٩٠٨	٦

$$\text{الاحداثي الرأسي المتوسط} = \frac{١ \times ١٠٠٠}{٢٫٦٢٦} \times ٠٫٣٩٨٩٤ = ١٥١٫٩$$

وتكون الخطوة التالية إجراء اختبار كاي تربيع للحكم على حسن المطابقة بين التكرارات الواقعية والتكرارات النظرية التي استنتجناها من عملية التوفيق .

لماذا نوفق المنحنى المعتدل :

يمثل التوزيع التكراري عادة عينة مسحوبة من مجتمع معين على أساس ان تمثل صفات هذا المجتمع الذي انتخبت منه العينة ، وحيث انه لا يمكن دراسة جميع وحدات المجتمع الذي تمثله العينة فاننا نستنتج هذه الصفات من دراستنا للعينة ، ولهذا نقوم بعملية توفيق المنحنى المعتدل أو أي منحنى آخر محاولين أن نعرف كيف يكون شكل المنحنى الذي يمثل المجتمع كله نظراً لأن المنحنى الذي يمثل العينة قد يكون به بعض التعرجات الناشئة عن ظهور بعض الوحدات الشاذة بوزن يفوق وزنها في المجتمع . ونحن عندما نوفق المنحنى المعتدل لتوزيع خاص بعينة ما فاننا بذلك نبدي رأياً، وطبعاً لا يكون رأياً اعتباطياً وإنما يكون قائماً على اساس ما نشاهده من بعض التماثل في التوزيع خاصة اذا كان معامل الالتواء الذي نحسبه صغيراً جداً . ولا يجب ان نتوقف عند إبداء الرأي فقط بل نقوم بعملية التوفيق حتى نستنتج ما نتوقعه من تكرارات كي يمكن ان نجري عليها اختبار حسن المطابقة (اختبار كاي تربيع) الذي يمكننا من تأكيد الرأي او استبعاده .

كذلك نستطيع بتوفيق المنحنى المعتدل ان نحدد تعميما على المجتمع الذي تمثله العينة نسبة الوحدات التي نتوقع ان تكون أقل أو أكثر أو بين قيمتين معينتين. مثلاً اذا وفقنا المنحنى الذي يمثل طول عمر اللمبات الكهربائية ان نعرف تعميما على مجموع اللمبات الكهربائية المنتجة نسبة اللمبات التي سوف تستمر في

الاضاءة ١٥٠٠ ساعة أو أكثر أو أقل من هذا الزمن او غيره . كذلك يمكن بتوفيق المنحنى الخاص بتوزيع يمثل الوفيات حسب فئات العمران نعرف تعميما نسبة عدد الافراد الذين يتوفون في كل فئة عمر . كذلك نستطيع بتوفيق المنحنى المعتدل ان نحدد على اساس البيانات التي نأخذها من عينة ما خاصة بمقاييس الرجال او النساء او الاطفال نسبة الوحدات التي يمكن انتاجها من كل نوع من الملابس تبعاً للمقاييس المختلفة التي نريدها .

استعمال التوزيع المعتدل في اختبار الفروض .

على اساس توزيع ظاهرة ما توزيعاً معتدلا نستطيع ان نحدد ما اذا كانت وحدة ما بالنسبة لهذه الظاهرة هي وحدة عادية او غير عادية (شاذة) وفي الغالب تكون الوحدة عادية اذا كانت ضمن الـ ٩٥ ٪ ، ومشكوكا في امرها اذا كانت واقعة بين ٩٥ ٪ و٩٩ ٪ ، وشاذة اذا كانت خارج الـ ٩٩ ٪ ومن الواضح ان تحديد وقوع الوحدة في اي قطاع من هذه القطاعات يتوقف على بعد قيمتها عن الوسط الحسابي محسوبة بالدرجات المعيارية اي نسبة الفرق الى الانحراف المعياري للتوزيع .

مثال ٢٩ :

اذا كان متوسط اجر العامل في احدى الصناعات ٣٥ ليرة في الاسبوع ، بانحراف معياري ٩ ليرات . ووجد ان عامل ما يتقاضى ٦٠ ليرة في الاسبوع هل اجر هذا العامل يختلف عن اجور عمال هذه الصناعة .

الدرجة المعيارية = $\frac{٦٠ - ٣٥}{٩}$ ، $\frac{٣٥}{٩}$: ٢٫٧٨ وهو اكبر من مستوى معنوية ٢٫٥٨ المقابل لدرجة ثقة ٩٩ ٪ . وبذلك نرفض الفرض بان هذا العامل لا يختلف اجره عن اجور العمال في الصناعة . أي ان هناك اختلاف

جوهري بين اجره والاجور عامة في هذه الصناعة ويمكن الاجابة على السؤال بطريقة اخرى .

حدي الفئة بمستوى معنوية ٩٩ ٪ = ٣٥ ± ٢٫٥٨ × ٩

وبمستوى معنوية ٩٥ ٪ = ٣٥ ± ١٫٩٦ × ٩

ونلاحظ ان الاجر ٦٠ يقع خارج هذا المدى ، وبذلك يكون هناك فرق جوهري بين اجر العامل والاجور عامة في هذه الصناعة .

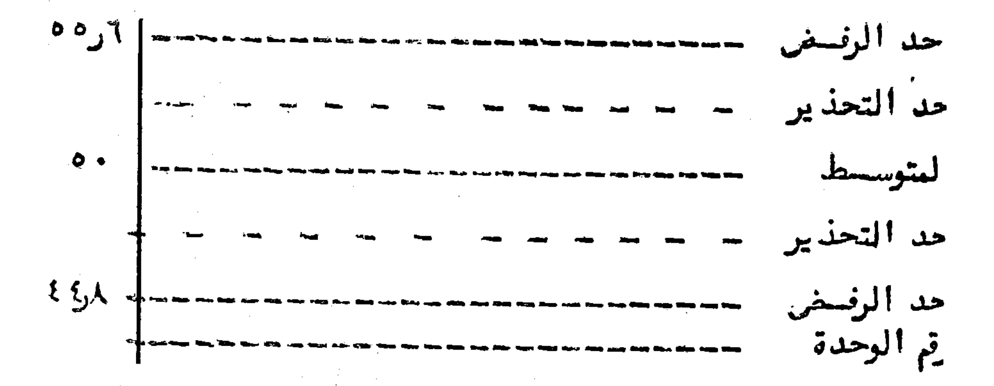
وعلى اساس هذه الفائدة للتوزيع المعتدل يمكن ضبط جودة الانتاج في مصنع ما حيث نستطيع أن نتعرف على الوحدات التي يكون انتاجها شاذاً بالنسبة للمتوسط العام الخاص بانتاج هذا المصنع .

مثال ٣٠ – يقوم أحد المصانع بانتاج الاسمنت وقد ضبطت آلات التعبئة بحيث تملأ ٥٠ كلـــغ في كل كيس بانحراف معياري ٢ كجم والمطلوب رسم حدود ضبط الانتاج .

بمستوى معنوية ٩٥٪ يكون حدي الثقة ٥٠ ± ١٫٩٦ × ٢
= ٥٣٫٩٢ كجم ، ٤٦٫١٨ كجم

وبمستوى معنوية ٩٩٪ يكون حدي الثقة ٥٠ ± ٢٫٥٨ × ٢
= ٥٥٫١٦ كجم ، ٤٤٫٨٤ كجم

وبذلك يمكن رسم خريطة لمراقبة الانتاج على أساس هذه الحدود للثقـة كالآتي : –



فاذا عبئت الآلة الأكياس بالاوزان الآتية ٥١ ، ٥٢٫٦ ، ٥٠ ، ٥٤٫٩ ، ٥١٫٨ ، ٥٧ ، ٤٩٫٥ كجم ، نلاحظ أن السابعة مرفوضة والثانية والخامسة مشكوك في امرهما وباقي الأكياس مقبولة .

ومن الواضح اننا في الاختبار السابق تعتبر الوحدة شاذة اذا كانت مفرطة في الزيادة أو مفرطة في النقص . فاذا كنا نريد أن نجيب على السؤال ما اذا كانت الوحدة مفرطة في الزيادة فقط أو مفرطة في النقصان فقط ، فاننا في هذه الحالة نهتم بناحية المنحنى التي تتعلق بالاختبار ، فاذا كنا نسأل هل الوحدة مفرطة في الزيادة لا يهمنا الجانب الأيسر من المنحنى ، واذا كنا نسأل هل الوحدة مفرطة في النقصان لا يهمنا الجانب الايمن من المنحنى . فاذا كنا نهتم بالجانب الايمن نضيف ٢٫٥٨ع الى الوسط الحسابي ، واذا كان اهتمامنا موجهاً الى الجانب الايسر في المنحنى نطرح ٢٫٥٨ع من الوسط الحسابي :

تمـــارين

١ – وفق المنحنى المعتدل للتوزيع الآتي بطريقتين :

ف	ك
٨٠–	١
٨١–	٨
٨٢–	٣٥
٨٣–	٨٢
٨٤–	١٢٢
٨٥–	١٢٤
٨٦–	٨٣
٨٧–	٣٤
٨٨–	١٠
٨٩– ٩٠	١

(التكرارات بعد التوفيق هي ٢ ، ٩ ، ٣٤ ، ٨١ ، ١٢٤ ، ١٢٤ ، ٨١ ، ٣٤ ، ٩ ، ٢)

٢ – في توزيع طبيعي ٧٪ من عدد القيم أقل من ٣٥ ، ٨٩٪ من هذه القيم أقل من ٦٣. ما هو الوسط الحسابي والانحراف المعياري لهذا التوزيع .

سَ = ٥٠٫٣   ع = ١٠٫٣٩

٣ – اذا كان متوسط درجات الطلبة ٧٢ بانحراف معياري ٩ – فـاذا تقرر أن يعطى ١٠٪ من الطلبة درجة ممتاز ، ما هي أقل درجة يجب أن يحصل عليها هؤلاء الطلبة (الدرجة = ٨٤) .

٤ – اذا كان احد المصانع ينتج انابيب متوسط قطرها ٠٫٦١٤٠ بوصة بانحراف معياري ٠٫٠٠٢٥ بوصة . احسب نسبة الانابيب التي ينتجها المصنع بالمقاييس الآتية :

أ – بين	٠٫٦١	بوصة	و٠٫٦١٨	=	٩٣	٪
ب – اكبر من	٠٫٦١٧	بوصة		=	٨٫١	٪
ح – أقل من	٠٫٦٠٨	بوصة		=	٤٧	٪
و – قطرهايساوي	٠٫٦١٥	بوصة		=	١٥	٪

٥ – ينتج احد المصانع انابيب متوسط قطرها ٠٫٥٠٢ بوصة بانحراف معياري ٠٫٠٠٥ بوصة فاذا كانت الانابيب تقبل اذا كان قطرها يتراوح بين ٠٫٤٩٦ بوصة و٠٫٥٠٨ بوصة احسب نسبة الانابيب المرفوضة .

(النسبة = ٢٣ ٪)

ارسم لوحة لمراقبة الانتاج في هذا المصنع وفسر كيف يمكن استعمالها .

٦ – اذا كان متوسط الاوزان لخمسمائة طالب هو ١٥١ رطل بانحراف معياري ١٥ رطل كم عدد الطلبة الذين يتراوح اوزانهم بين ١٢٠–١٥٥ رطل (٣٠٠ طالب) . وكم عدد الطلبة الذين تزيـــد اوزانهم عن ١٨٥ رطل (٥ طلاب) .

الفصل التاسع

اختبار كاي تربيع (χ^2)

يواجه الاحصائي مشكلة تتخذ صوراً مختلفة في كثير من الدراسات الاحصائية ، مشكلة اختبار التطابق بين بيانات واقعية وبيانات مستنبطة على اساس افتراض معين أو على أساس اجراء الدراسة بالمعاينه حيث نحصل على بيانات من العينة تختلف بعض الشيء عن المتوقع طبقاً لقوانين الاحتمالات، حتى يمكن بذلك ان يحكم على افتراضه فاما أن يطمئن اليه ويعتبره مناسباً غير بعيد عن الحقيقة واما أن لا يعتبره كذلك فيستبعده . واختبار كاي تربيع هو الاداة الاحصائية التي يمكن بواسطتها مواجهة هذه المشكلة والبت فيها . ولمناقشة طريقة اجراء هذا الاختبار يكون من الأفضل أن أورد أمثلة ملموسة توضح خطوات العمل اللازمة .

مثال ٣٠ :

في عام ١٩٥٠ قام مركز الدراسات الاستقصائية باحدى الجامعات الاميركية بدراسة حول العلاقة بين الحالة الزواجية لرب الاسرة والادخار ؛ ولقد تبين من المعلومات التي جمعت عن ٣٣٢٧ أسرة الوضع التالي فيما يختص بموضوع الدراسة :

الحالة الزواجية لرب الاسرة	عدد المدخرين	عدد غير المدخرين	المجموع
اعزب	٤٩٠	٣٩٠	٨٨٠
متزوج	١٥٥٢	٧٩٥	٢٤٤٧
المجموع	٢٠٤٢	١١٨٥	٣٣٢٧

من هذه المعلومات يرغب الاحصائي في الاجابة على السؤال الذي يتبادر الى الذهن في هذه الحالة وهو هل هناك اختلاف جوهري بين سلوك العزاب والمتزوجين نحـــو الادخار او ان الاختلاف بين سلوكهم ليس الا اختلاف ظاهري يمكن ان يكون راجعاً لمجرد الصدفة الناتجة من دراسةعينة من الاسر جاءت نتائجها معطية بعض الوزن (عشوائياً) لاحـدى الفئات الاربع التي يتكون منها الجدول السابق

للاجابة على هذا السؤال يبدأ الاحصائي بافتراض ان لا فرق هناك بين سلوك كل من الفريقين ، أي انهما يكونا مجتمعاً واحداً سلوكه متشابه نحو الادخار . وبمعنى آخر يفترض الاحصائيأن لا علاقة البتة بين الحالةالزواجية لرب الاسرة وسلوكه نحو الادخار اي لا توجد أية علاقة بين تصنيف أرباب الاسر الى مدخرين وغير مدخرين وتصنيفهم الى عزاب ومتزوجين . علىهذا الأساس تكون النسبة العامة للمدخرين في هذا المجتمع المفترض هي $\frac{٣٢٤٢}{٣٣٢٧}$

(مجموع المدخرين في العينة الى المجموع الكلي لوحداتها)

على اساس هذه النسبة العامة وتبعاً للافتراض بأن كل من الفريقين ليس الا صورة ينعكس عليها السلوك العام في المجتمع انعكاساً كاملا يكون عدد المدخرين وغير المدخرين في كل من الفريقين كالآتي :

العزاب:

$$٨٨٠ \times \frac{٢٠٤٢}{٣٣٢٧} =$$ او ٥٤٠ مدخرين

٨٨٠ - ٥٤٠ر١ = ٣٣٩ر٩ غير مدخرين

المتزوجون:

$$٢٤٤٧ \times \frac{٢٠٤٢}{٣٣٢٧} =$$ ١٥٠١ر٢ مدخرين

٢٤٤٧ – ١٥٠١ر٩ غير مدخرين

لما كانت الارقام المحسوبة للمدخرين وغير المدخرين في كل من الفريقين قد أجري حسابها على أساس الافتراض الذي بدأنا به ، لذا نسميها بالتكرارات النظرية أي الأرقام التي كان يمكن ان تظهر في الجدول السابق لو كانت الحالة الزواجية لرب الأسرة وسلوكه نحو الادخار ظاهرتان مستقلتان عن بعضهما اي لا علاقة البتة بينهما . وباستعراض هذه التكرارات النظرية وبمقارنتها مع التكرارات الواقعية الواردة في الجدول الأصلي نلاحظ وجود فرق بينهما ، والمشكلة التي تواجهنا بعد ذلك هي كيف يمكن ان نحكم على هذا الفرق لنقرر ما إذا كان فرقاً بسيطاً يمكن أن يكون راجعاً إلى الصدفة أو انه من الكبر بحيث لا يمكن ان يكون سببه مجرد الصدفة . ولا شك اننا بحكمنا هذا نستطيع أن نقرر ما إذا كان افتراضنا الذي بدأنا بـه افتراض سليم أو غير سليم ، ذلك لأن الوضع الذي يمكن أن نصل اليه هو احد أمرين ، اما انه على أساس افتراضنا استنبطنا بيانات لا تختلف عن البيـــانات الواقعية الا اختلافاً بسيطاً وبذلك يكون هناك احتمال كبير في صحة الافتراض واما العكس إذا اختلفت البيانات المستنبطة عن البيانات الواقعية اختلافاً جوهريا وبذلك يكون هناك احتمال كبير في عدم صحة الافتراض ، الأمر الذي يؤدي الى ضرورة رفضه واستبعاده .

ان بعض الأرقام النظرية يمكن ان تكون أقل بينما يمكن ان يكون بعضها أكثر من الأرقام الواقعية وبـذلك فان جمع هذه الفروق بإشاراتها الجبرية يؤدي الى نتيجة مضللة في الحـكم على المجموع الكـــلي للفروق بين التكرارات الواقعية والنظرية . ونظراً الى ان اهتمامنا ليس موجهاً الى كون الفرق بالزائد أو بالناقص وإنما لقيمة الفرق نفسه ، لذا وحتى نتخلص من من الاشارات الجبرية نربع الفروق . ولما كانت قيمة الفرق لا يمكن ان تعطينا فكرة سليمة من معنويته (اذا كان فرقاً بسيطاً أو كبيراً) حيث يمكن أن يكون فرقاً ما – واحد مثلا – فرقاً كبيراً إذا كان نتيجة طرح أرقام صغيرة جداً ويمكن أن يكون فرقاً صغيراً جدا اذا كان نتيجة طرح أرقام

كبيرة جداً ، وبذلك حتى يمكن أن نأخذ فكرة سليمة عن معنوية الفرق ننسبه الى التكرارات النظرية التي حسبناها وفيما يلي هذه العمليات الحسابية منظمة في شكل جدول :

التكرارات الواقعية	التكرارات النظرية	الفرق	مربع الفرق	مربع الفرق / التكرارات النظرية
٤٩٠	٥٤٠ر١	٥٠ر١	٢٥١٠ر٠١	٤ر٦٤٧٣
١٥٥٢	١٥٠١ر٩	٥٠ر١	٢٥١٠ر٠١	١ر٦٧١٢
٣٩٠	٣٣٩ر٩	٥٠ر١	٢٥١٠ر٠١	٧ر٣٨٤٦
٨٩٥	٩٤٥ر١	٥٠ر١	٢٥١٠ر٠١	٢ر٦٥٥٨
			المجموع	١٦ر٣٥٨٩

هذه القيمة ١٦ر٣٥٨٩ هي التي يطلق عليها اصطلاح كاي تربيع وهي التي تريد أن تحكم على معنويتها ، أي على مـــا اذا كانت قيمة جوهرية أو قيمة بسيطة يمكن أن تكون راجعة الى الصدفة . وحتى نستطيع ذلك يجب أن نرجع الى جدول كاي تربيع ، وهو جدول توزيع القيم المختلفة لكاى تربيع عند ما تكون الصدفة هي العامل الوحيد الذي يلعب دوره في ايجاد الفروق بين التكرارات المشاهدة (الواقعية) والتكرارات (النظرية) ، وقد اعد هذا الجدول كارل بيرسون ليكون اداة موضوعية (غـير خاضعة للتقدير الشخصي للباحث) يمكن على أساسها الحكم على معنوية الفروق وتحديد احتمال الصدفة في الحصول عليها .

لا شك أن معنوية كاى تربيع (١٦ر٣٥٨٩ في مثالنا هذا) تتوقف على عدد القيم التي جمعت حيث يمكن أن تكون قيمة ذات معنى اذا كانت مجموعاً لعدد

بسيط جداً من القيم ويمكن أن تكون قيمة بسيطة اذا كانت مجموعاً لعدد كبير جداً من القيم . لهذا وحتى نستطيع أن نحكم على معنوية كاى تربيع حكماً دقيقاً يجب أن نحدد عدد القيم التي جمعت للتوصل الى هذه القيمة (كاى تربيع). الا اننا نلاحظ اننا في تطبيقنا للافتراض الذي بدأنا به حتى نتوصل الى التكرارات النظرية لم نكن احرار مطلق الحرية في ذلك حيث كنا دائماً مقيدين بمجموع كل فئة من الفئات (مجموع المدخرين ومجموع غير المدخرين ثم مجموع العزاب ومجموع المتزوجين)، اذ لا يجب أن نغير من هذه المجاميع عند تطبيق الافتراض، وبمعنى آخر كانت حريتنا قاصرة على توزيع هذه المجاميع بين الفئات المختلفة ، لهذا كان يكفي أن نحسب التكرار النظري لاحدى هذه الفئات الاربع على أساس الافتراض اما باقي التكرارات النظرية فعند حسابها كنا مقيدين بالمجامع الواقعية ، وبذلك فان معنوية كاى تربيع مرتبطة بدرجة حرية واحدة وليست باربع درجات ، وبمعنى آخر هناك تكرار واحد فقط الذي علىأساسه يمكن أن نحكم على معنوية كاى تربيع وليس اربع تكرارات . هذا هو الاساس النظري لما نسميه بدرجات الحرية التي تظهر في كعب الجدول الذي أعده بيرسون فيما يختص بتوزيع كاى تربيع . ويمكن حساب عدد دوجات الحرية حساباً سريعاً بطرح واحد من عدد الفئات الافقية التي يتكون منها الجدول وضرب الناتج في حاصل طرح واحد من عدد الفئات الرأسية التي تظهر في الجدول .

أي (٢ — ١) (٢ — ١) = ١ في مثالنا السابق .

أو بطرح عدد الثوابت التي استخدمت في تقدير التكرارات النظرية من عدد المشاهدات (التكرارات النظرية) أي ٤ — ٣ = ١

وبالرجوع الى جدول كاى تربيع وأمام درجة حرية واحدة نجد الوضع التالي : —

درجات الحرية	احتمال الحصول على قيم كاي تربيع بطريقة المصادفة						
	٠٠١ ر	٠١ ر	٠٥ ر	١٠ ر	٥٠ ر	٩٠ ر	٩٩ر
١	١٠ ر٨٢٧	٦ر٦٢٥	٣ر٨٤١	٧٠٦ر	٤٥٥ر	٠ر ١٥٨	١٥٧ر

ولما كانت قيمة كاي تربيع في مثالنا السابق هي ١٦ر٣٥٨٩ فاحتمال الحصول على هذه القيمة بطريقة المصادفة احتمال صغير جداً أقل من ٠٠١ر أي أقل من ١ر ٪ .

والمشكلة التي تواجهنا الآن هي ما هو الاحتمال الذي يمكن أن نعتبره صغيراً وذلك الذي يمكن أن نعتبره احتمالاً كبيراً ؟ لقد جرى العرف في الدراسات الاجتماعية على ان احتمال الصدفة ٥ ٪ أو أقل يعتبر احتمالاً صغيراً، وانه إذا زاد احتمال الصدفة عن ٥ ٪ يعتبر احتمالاً كبيراً .

وبذلك فاحتمال ان نحصل على القيمة ١٦ر٣٥٨٩ بطريق المصادفة احتمال صغير جداً أي ان هذه القيمة تمثل فرقاً جوهرياً بين التكرارات الواقعية (المشاهدة) والتكرات النظرية المتوقعة على أساس الافتراض الذي بدأنا به . وبمعنى آخر يمكن أن نقول انه على أساس افتراضنا بالتشابه بين سلوك العزاب والمتزوجين نحو الادخار أي انهم يكونوا مجتمعاً واحداً توقعنا بيانات اختلفت اختلافاً جوهريا عن البيانات الواقعية وبهذا نستطيع أن نقول انه على أساس اختبارنا هذا هناك احتمال كبير في عدم صحة هذا الافتراض . لذلك فان القول بعدم وجود أي علاقة بين الحالة الزواجية وبين السلوك نحو الادخار قول يجب استبعاده حيث ان البيانات المتوفرة لدينا تقيم الدليل القوي على وجود العلاقة بين هاتين الظاهرتين ، وبمعنى آخر تدل البيانات المشاهدة (الواقعية) على ضعف الادخار بين العزاب

وشيوعه بين المتزوجين بـدرجة تزيد عما نتوقع على أساس الافتراض الذي بدأنا به .

مثال ٣١ :

قام باحث في استهلاك نوع من الحلوى بتقسيم دولة ما إلى ثمان مناطق وأخـذ من كـل منطقة عينة عشوائية من السكان صنفهم تبعاً لاستهلاكهم وعدم استهلاكهم لهذا النوع ، وكانت البيانات التي حصل عليها كالآتي :

رقم المنطقة	عدد المستهلكين	عدد غير المستهلكين
١	٥٦	١٧
٢	٨٧	٢٠
٣	١٤٢	٥٨
٤	٧١	٢٠
٥	٧٧	٣١
٦	٧٢	٢٣
٧	١٠٠	٢٥
٨	١٤٢	٣٨
المجموع	٧٥٨	٢٤٢

ويريد الباحث أن يعرف ما إذا كان استهلاك هذا النوع من الحلوى يختلف اختلافاً جوهرياً بين الثمان مناطق التي تتكون منها الدولة أو ان الاختلاف ليس الا اختلافاً بسيطاً يمكن أن نعزوه الى عشوائية الدراسة بالمعـاينة .

وللاجابة على هـ ا الاستفسار نبدأ بافتراض ان استهلاك هذا النوع من الحلوى لا يختلف بين هذه المناطق وانها جميعاً تكون مجتمعاً واحداً تكون فية النسبة العامة لمستهلكى هـــذا النوع هي $\frac{٧٥٨}{١٠٠٠}$.

وبذلك فكل منطقة يمكن أن نعتبرها صورة صحيحة ينعكس عليها هذا السلوك العام نحو استهلاك هذا النوع من الحلوى – وعلى هذا الأساس يمكن أن نحسب عدد المستهلكين وغـــير المستهلكين في كل منطقة كالآتي (التكرارات النظرية) :

المنطقة ١ – عـــدد المستهلكين ٧٣ × $\frac{٧٥٨}{١٠٠٠}$ = ٥٥ تقريباً

وعدد غير المستهلكين = ٧٣ – ٥٥ = ١٨ .

المنطقة ٢ – عـــدد المستهلكين ١٠٧ × $\frac{٧٥٨}{١٠٠٠}$ = ٨١ تقريباً

وعدد غير المستهلكين = ١٠٧ – ٨١ = ٢٦ .

المنطقة ٣ – عـــدد المستهلكين ٢٠٠ × $\frac{٧٥٨}{١٠٠٠}$ = ١٥٢ تقريباً

وعدد غير المستهلكين = ٢٠٠ – ١٥٢ = ٤٨ .

المنطقة ٤ – عـــدد المستهلكين ٠١ × $\frac{٧٥٨}{١٠٠٠}$ = ٦٩ تقريباً

وعدد غير المستهلكين = ٩١ – ٦٩ = ٢٢ .

المنطقة ٥ – عـــدد المستهلكين ١١٩ × $\frac{٧٥٨}{١٠٠٠}$ = ٩٠ تقريباً

وعدد غير المستهلكين = ١١٩ – ٩٠ = ٢٩

المنطقة ٦ – عــــدد المستهلكين ٩٥ $\times \frac{٧٥٨}{١٠٠٠}$ = ٧٢ تقريباً

وعدد غير المستهلكين = ٩٥ – ٧٢ = ٢٣ .

المنطقة ٧ – عــــدد المستهلكين ١٢٥ $\times \frac{٧٥٨}{١٠٠٠}$ = ٩٥ تقريباً

وعدد غير المستهلكين = ١٢٥ – ٩٥ = ٣٠ .

المنطقة ٨ – عــــدد المستهلكين ١٩٠ $\times \frac{٧٥٨}{١٠٠٠}$ = ١٤٤ تقريباً

وعدد غير المستهلكين = ١٩٠ – ١٤٤ = ٤٦ .

والخطوة التالية هي تحديد الفرق (كاي تربيع) بين التكرارات الواقعية والتكرارات النظرية :

المنطقة	التكرارات الواقعية (المشاهدة)	التكرارات النظرية (المتوقعة)	ف	ف٢	ف٢ / التكرارات النظرية
١	٥٦	٥٥	١	١	ر٠١٨
	١٧	١٨	١	١	ر٠٥٥
٢	٨٧	٨١	٦	٣٦	ر٤٤٥
	٢٠	٢٦	٦	٣٦	١ر٣٨٤
٣	١٤٢	١٥٢	١٠	١٠٠	ر٦٥٧
	٥٨	٤٨	١٠	١٠٠	٢ر٠٨٣
٤	٧١	٦٩	٢	٤	ر٠٥٧
	٢٠	٢٢	٢	٤	ر١٨١
٥	٨٨	٩٠	٢	٤	ر٠٤٥
	٣١	٢٩	٢	٤	ر١٣٧
٦	٧٢	٧٢	صفر	صفر	صفر
	٢٣	٢٣	صفر	صفر	صفر
٧	١٠٠	٩٥	٥	٢٥	ر٢٦٣
	٢٥	٣٠	٥	٢٥	ر٨٣٤
٨	١٤٢	١٤٤	٢	٤	ر٠٢٧
	٤٨	٤٦	٢	٤	ر٠٨٧
المجموع					٦ر٢٧٣

والخطوة التالية تحديد عدد درجات الحرية وهي في هذا المثال (٨ — ١) (٢ — ١) = ٧ ، وبشكل آخر ١٦ تكرار نظري ناقص عدد الثوابت التي استخدمت في حساب هذه التكرارات النظرية ، وهي في هذه الحالة ٩ ثوابت — أي ١٦ — ٩ = ٧ درجات حرية .

من جدول كاي تربيع نحصل على البيانات التالية :

درجات الحرية	احتمال الحصول على قيمة كاي تربيع بطريق المصادفة						
	٠٠١ر	٠١ر	٠٥ر	١٠ر	٥٠ر	٩٠ر	٩٩ر
٧	٢٤ر٣٢	١٨ر٤٧٥	١٣٠٦٧	١٢ر٠١٧	٦ر٣٤٦	٢ر٨٣٣	٢٣٩ر

وبذلك فاحتمال الحصول على القيمة ٦ر٢٧٣ بطريق المصادفة اكثر من ٥٠٪ وهو احتمال كبير جداً ، وبذلك ليس هناك ما يجعلنا نشك في افتراضنا بان استهلاك هذا النوع من الحلوى متشابه بين المناطق المختلفة التي تتكون منها الدولة . وبشكل آخر يمكننا أن نقول ان هذا الاختبار لا يقيم الدليل ضد الفرض الذي وضعناه . وتفسير ذلك اننا على أساس افتراضنا بان البيانات المتوفرة لدينا (الواقعية) لا توحي بالاختلاف بين هذه المناطق بالنسبة الى استهلاك هذا النوع من الحلوى استنتجنا بيانات (التكرارات النظرية) لم تختلف عن البيانات الواقعية إلا اختلافاً بسيطاً احتمال أن يكون راجعاً إلى مجرد الصدفة احتمال كبير .

مثال ٣٢ :

قام و. ف. ولدون بتجربة رمي ١٢ زهرة نرد دفعة واحد وقد كرر هذه التجربة ٤٠٩٦ مرة ولاحظ في كل مرة عدد الزهرات الناجحة (النجاح في هذه التجربة هو ظهور ٦ أو ٥ أو ٤) فحصل على التوزيع الآتي :

عدد الزهرات الناجحة	التكرار المشاهد
صفر	صفر
١	٧
٢	٦٠
٣	١٩٨
٤	٤٣٠
٥	٧٣١
٦	٩٤٨
٧	٨٤٧
٨	٥٣٦
٩	٢٥٧
١٠	٧١
١١	١١
١٢	صفر

المجموع ٤٠٩٦

في هذا المثال نلاحظ ان احتمال النجاح بالنسبـــة للزهرة الواحدة = $\frac{٣}{٦}$ = ١ / ٢ واحتمال الفشل = $\frac{٣}{٦}$ = ١ / ٢ كذلك، ورمي ١٢ زهرةدفعة واحدة يعني عمل ١٢ محاولة في كل رمية . وعلى هذا الأساس يمكننا أن نحدد عدد المرات من (ال ٤٠٩٦ مرة) التي يمكن أن لا ينجح فيها أية زهرة وتلك التي يمكن أن ينجح فيها زهرة واحـدة ثم التي تنجح فيهـا زهرتين وهكذا .

عدد المرات التي لا تنجح فيها زهرة واحدة =

$$١٢\ ق_{٠} \times \left(\frac{١}{٢}\right)^{\text{صفر}} \times \left(\frac{١}{٢}\right)^{١٢} \times ٤٠٩٦ = ١ \text{ مرة}$$

عدد المرات التي تنجح فيها زهرة واحدة =

$$١٢\ ق_{١} \times \frac{١}{٢} \times \left(\frac{١}{٢}\right)^{١١} \times ٤٠٩٦ = ١٢ \text{ مرة}$$

عدد المرات التي تنجح فيهـــا زهرتان =

$$١٢\ ق_{٢} \times \left(\frac{١}{٢}\right)^{٢} \times \left(\frac{١}{٢}\right)^{١٠} \times ٤٠٩٦ = ٦٦ \text{ مرة}$$

عدد المرات التي تنجح فيها ثلاث زهرات =

$$١٢\ ق_{٣} \times \left(\frac{١}{٢}\right)^{٣} \times \left(\frac{١}{٢}\right)^{٩} \times ٤٠٩٦ = ٢٢٠ \text{ مرة}$$

عدد المرات التي تنجح فيها أربع زهرات =

$$١٢\ ق_{٤} \times \left(\frac{١}{٢}\right)^{٤} \times \left(\frac{١}{٢}\right)^{٨} \times ٤٠٩٦ = ٤٩٥ \text{ مرة}$$

عدد المرات التي تنجح فيها خمس زهرات =

$$١٢\ ق_{٥} \times \left(\frac{١}{٢}\right)^{٥} \times \left(\frac{١}{٢}\right)^{٧} \times ٤٠٩٦ = ٧٩٢ \text{ مرة}$$

عدد المرات التي تنجح فيها ست زهرات =

$$١٢\ ق_{٦} \times \left(\frac{١}{٦}\right)^{٦} \times \left(\frac{١}{٢}\right)^{٦} \times ٤٠٩٦ = ٩٢٤ \text{ مرة}$$

عدد المرات التي تنجح فيها سبع زهرات =

$$١٢ ق_{٧} \times \left(\frac{١}{٦}\right)^{٧} \times \left(\frac{١}{٦}\right)^{٥} \times ٤٠٩٦ = ٧٩٢ \text{ مرة}$$

عدد المرات التي تنجح فيها ثمان زهرات =

$$١٢ ق_{٨} \times \left(\frac{١}{٦}\right)^{٨} \times \left(\frac{١}{٦}\right)^{٤} \times ٤٠٩٦ = ٤٩٥ \text{ مرة}$$

عدد المرات التي تنجح فيها تسع زهرات =

$$١٢ ق_{٩} \times \left(\frac{١}{٦}\right)^{٩} \times \left(\frac{١}{٦}\right)^{٣} \times ٤٠٩٦ = ٢٢٠ \text{ مرة}$$

عدد المرات التي تنجح فيها عشر زهرات =

$$١٢ ق_{١٠} \times \left(\frac{١}{٦}\right)^{١٠} \times \left(\frac{١}{٦}\right)^{٢} \times ٤٠٩٦ = ٦٦ \text{ مرة}$$

عدد المرات التي تنجح فيها احد عشر زهرة =

$$١٢ ق_{١١} \times \left(\frac{١}{٦}\right)^{١١} \times \frac{١}{٦} \times ٤٠٩٦ = ١٢ \text{ مرة}$$

عدد المرات التي تنجح فيها اثني عشر زهرة =

$$١٢ ق_{١٢} \times \left(\frac{١}{٦}\right)^{١ر٢} \times \left(\frac{١}{٦}\right)^{\text{صفر}} \times ٤٠٩٦ = ١ \text{ مرة}$$

يتبين لنا أن هنالا فرق بين عدد مرات النجاح المتوقعة وعـــدد مرات النجاح المشاهدة فعلاً ونريد الآن أن نختبر الفرق بينهما حتى نحدد ما اذا كان بسيطاً يمكن أن نعزوه الى مجرد المصادفة أو أنـه فرق جوهري يمكن على

أساسه أن نشك في عدم تحيز زهر النرد الذي نقوم برميه في هذه التجربة . وللاجابة على ذلك نقوم باجراء اختبار كاي تربيع كالآتي :

عدد الزهرات الناجحة	التكرارات المشاهدة	التكرارات المتوقعة	ف	ف٢	ف٢ / التكرار المتوقع
صفر	صفر	١	– ١	١	١ر٠٠٠٠
١	٧	١٢	– ٥	٢٥	٢ر٠٨٣٣
٢	٦٠	٦٦	– ٦	٣٦	٠ر٥٤٥٥
٣	١٩٨	٢٢٠	– ٢٢	٤٨٤	٢ر٢٠٠٠
٤	١٣٠	٤٩٥	– ٦٥	٤٢٢٥	٨ر٥٣٥٤
٥	٧٣١	٧٩٢	– ٦١	٣٧٢١	٤ر٦٩٨٢
٦	٩٤٨	٩٢٤	+ ٢٤	٥٧٦	٠ر٦٣٣٤
٧	٨٤٧	٧٩٢	+ ٥٥	٣٠٢٥	٣ر٨١٩٤
٨	٥٣٦	٤٩٥	+ ٤١	١٦٨١	٣ر٣٩٦٠
٩	٢٥٧	٢٢٠	+ ٣٧	١٣٦٩	٦ر٢٢٤٧
١٠	٧١	٦٦	+ ٥	٢٥	ر٣٧٨٨
١١	١١	١٢	– ٢	٤	ر٣٠٧٧
١٢	صفر	١			
			المجموع		٣٣ر٨١٠٤

وفي هذا المثال يكون عدد درجات الحرية (١٣ – ١) = ١٢ درجة

حيث طرحنـــا درجة واحدة نظراً لأننا كنا مقيدين بالمجموع الكلي لعدد المحاولات (٤٠٩٦ محاولة) عند حساب التكرارات المتوقعة . وبالنظر الى جدول كاى تربيع نحصل على البيانات التالية :

درجات الحرية	احتمال الحصول على قيمة كاي تربيع بطريقة المصادفة						
	٠٠١ر	٠١٠ر	٠٥ر	١٠ر	٥٠ر	٩٠ر	٩٩ر
١٢	٣٢ر٩٠٩	٢٦ر٢١٧	١٢ر٠٢٦	١٨ر٥٤٩	١١ر٣٤٠	٦ر٣٠٤	٣ر٥٧١

وبذلك يكون احتمال الحصول على القيمة ٣٣ر٨١٠٤ (قيمة كاى تربيع في مثالنا هذا) بطريق المصادفة احتمال ضعيف جداً ، فهي لذلك تشكل فرقاً جوهريا بين التكرارات المشاهدة والتكرارات المتوقعة ، الأمر الذي يجعلنا نشك في عدم تحيز زهر النرد المستعمل في التجربة ، وبمعنى آخر فإن اختبار كان تربيع يوحي الينا بتحيز زهر النرد المستعمل في التجربة ، هذا اذا كان لدينا ثقة تامة في عشوائية القائم بالتجربة .

مثال ٣٣ :

في احدى التجارب التي قام بها مندل على استنبات بعض البذور حصل على ٣١٥ بذرة صفراء ومستديرة و ١٠١ بذرة صفراء ومجعدة و ١٠٨ بذرة خضراء ومستديرة و ٣٢ بذرة خضراء ومجعدة . هـــل تعتبر هذه النتيجة مطابقة للنظرية الوراثية الخاصة بهذا الموضوع حيث نتوقـــع ان توزع هذه البذور بين هذه الفئات الاربع بنسبة ٩ : ٣ : ٣ : ١ .

على أساس هذا التوزيع المتوقع كان يجب ان نحصل على :

٥٥٦ (مجموع البذور) × $\frac{٩}{١٦}$ = ٢١٣ تقريباً بذرة صفراء ومستديرة

٥٥٦ « × $\frac{٣}{١٦}$ = ١٠٤ « « « ومجعدة

٥٥٦ « × $\frac{٣}{١٦}$ = ١٠٤ « « خضراء ومستديرة

٥٥٦ « × $\frac{١}{١٦}$ ٣٥ « « « «

وبالخطوة التالية نختبر الفرق بين هذه التكرارات المتوقعة والتكرارات المشاهدة في التجربة .

النوع	التكرار المشاهد	التكرار المتوقع	ف	ف٢	ف٢ / التكرار المتوقع
١	٣١٥	٣١٣	٢	٤	٠١٢ر
٢	١٠١	١٠٤	٣	٩	٠٨٦ر
٣	١٠٨	١٠٤	٤	١٦	١٥٣ر
٤	٢٢	٣٥	٣	٩	٢٥٧ر
		المجموع			٥٠٨ر

ولما كانت درجات الحرية هي ثلاث درجات (٤ تكرارات متوقعة — ثابت واحد وهو مجموع البـــذور) فان احتمال الحصول على القيمة ٥٠٨ ر بطريقة المصادفة احتمال كبير جداً يزيد عن ٩٠ ٪ وبذلك يمكن ان تعتبر

هذه القيمة فرقاً بسيطاً ، وبمعنى آخر فإن التكرارات المشاهدة لا تختلف عن التكرارات المتوقعة الا اختلافاً بسيطاً يمكن أن يكون راجعاً الى المصادفة .

مثال ٣٤ :

تدل البيانات التالية على عدد المواليد في انجلترا وويلز عام ١٩٤١ مصنفة تبعاً لأشهر السنة . والمطلوب استخدام اختبار كاى تربيع لمعرفة ما اذا كانت هناك تغيرات موسمية في المواليد على أساس هذه البيانات :

الشهر	عدد المواليد	الشهر	عدد المواليد
ينـاير	٥٠١٥٩	يوليـو	٤٩٣٩٥
فبراير	٤٥٨٨٥	اغسطس	٥٠٤٤٣
مارس	٥٠٨١٩	سبتمبر	٥١٥٦٢
ابريل	٤٩٠٧٠	اكتوبر	٥٠٢٢٤
مـايو	٥٠٧٧١	نوفمبر	٤٧١٦٨
يونيو	٤٦٧٨٨	ديسمبر	٥٠٥٢٩
	المجمـوع		٥٩٢٨١٣

على أساس افتراض أن لا يوجد هناك تغيرات موسمية فان عدد المواليد اليومي يتوقع أن يكون ٥٩٢٨١٣ ÷ ٣٦٥ = ١٦٢٤ر١ مولود تقريباً وبذلك يكون عدد المواليد لكل شهر من الأشهر ذات الثلاثين يوماً ٤٨٧٢٤ مولود ولكل شهر من الأشهر ذات الواحد والثلاثين يوماً ٥٣٠٤٩ مـولود ولشهر فبراير ٤٥٤٧٦ مولود .

والخطوة التالية هي اختبار الفرق بين هذه البيانات المتوقعة والبيانات الواقعية كالآتي :

الشهر	عدد المواليد	العدد المتوقع	ف	ف٢	ف٢ / العدد المتوقع
يناير	٥٠١٥٩	٥٠٣٤٩	١٩٠	٣٦١٠٠	٠ر٧١٧
فبراير	٤٥٨٨٥	٤٥٤٧٦	٤٠٩	١٦٧٢٨١	٣ر٦٨٧
مارس	٥٠٨١٩	٥٠٣٤٩	٤٧٠	٢٢٠٩٠٠	٤ر٣٨٧
ابريل	٤٩٠٧٠	٤٨٧٢٤	٣٤٦	١١٩٧١٦	٢ر٤٥٧
مايو	٥٠٧٧١	٥٠٣٤٩	٤٢٢	١٨٧٠٨٤	٣ر٥٣٧
يونيو	٤٦٧٨٨	٤٨٧٢٤	١٩٣٦	٣٧٤٨٠٩٦	٧٦ر٨٢٥
يوليو	٤٩٣٩٥	٥٠٣٤٩	٩٥٤	٩١٠١١٦	١٨ر٠٧٦
اغسطس	٥٠٤٤٣	٥٠٣٤٩	٩٤	٨٨٣٦	ر١٧٥
سبتمبر	٥١٥٦٢	٤٨٧٢٤	٢٨٣٨	٨٠٥٤٢٤٤	١٦٥ر٣٠٣
اكتوبر	٥٠٢٢٤	٥٠٣٤٩	١٢٥	١٥٦٢٥	ر٣١٠
نوفمبر	٤٧١٦٨	٤٨٧٢٤	١٥٥٦	٢٤٢١١٣٦	٤٩ر٦٩١
ديسمبر	٥٠٥٢٩	٥٠٣٤٩	١٨٠	٣٢٤٠٠	ر٦٤٣
المجموع	٥٩٢٨١٣	٥٩٢٨١٥			٣٢٥ر٨٩٩

وعلى أساس ١١ درجة حرية (١٢ — ١) فإن احتمال الحصول على القيمة ٣٢٥ر٨٩٩ بطريقة المصادفة احتمال ضعيف جداً وبذلك فان هذا الاختبار يقيم الدليـــل على وجود تغيرات موسمية في المواليد حيث انه على اساس

افتراضنا بعدم وجود هذه التغيرات استنتجنا بيانات اختلفت عن الواقع اختلافاً جوهريا .

مثال ٣٥ :

البيانات الآتية تبين توزيع ٦٨٠٠ من الذكور في مدينة بادن تبعاً للون الشعر ولون العين كذلك :

لون العين	لون الشعر				المجموع
	اصفر	بني	اسود	احمر	
ازرق	١٧٦٨	٨٠٧	١٨٩	٤٧	٢٨١١
رمادي او اخضر	٩٤٦	١٣٨٧	٧٤٦	٥٣	٣١٣٢
بني او اسود	١١٥	٤٣٨	٢٨٨	١٦	٨٥٧
المجموع	٢٨٢٩	٢٦٣٢	١٢٢٣	١١٦	٦٨٠٠

والمطلوب دراسة ما اذا كانت هناك علاقة بين لون شعر الذكر ولون عينيه تبعاً لهذه البيانات . وبمعنى آخر هل هناك علاقة بين تصنيف هذه المجموعة من الذكور تبعاً للون الشعر ولون العينين او ان تصنيفهم تبعاً لهاتين الظاهرتين مستقل عن الآخر .

للاجابة على ذلك نبدأ بافتراض عدم وجود اية علاقة بين لون الشعر ولون العينين وان هذه المجموعة تكون مجتمعاً واحداً فيه نسبة ذوي العيون الزرق $\frac{٢٨١١}{٦٨٠٠}$ ونسبة ذوي العيون الخضر او الرمادية $\frac{٣١٣٢}{٦٨٠٠}$ ونسبة ذوي العيون السود $\frac{٨٥٧}{٦٨٠٠}$ وبتطبيق هذه النسب على المجموعات الاربع حسب لون الشعر نحصل

على البيانات النظرية الآتية :

مجموعة الرجال ذوي الشعر الأصفر:

٢٨٢٩ × ٢٨١١ / ٦٨٠٠ = ١١٦٩ ذوي عيونن زرق

٢٨٢٩ × ٣١٣٢ / ٦٨٠٠ = ١٣٠٣ ذوي عيون رمادية او خضر

٢٨٢٩ - (١١٦٩ + ١٣٠٣) = ٣٥٧ ذوي عيون بنية أو سود

مجموعة الرجال ذوي الشعر البني:

٢٦٣٢ × ٢٨١١ / ٦٨٠٠ = ١٠٨٨ ذوي عيون زرق

٢٦٣٢ × ٣١٣٢ / ٦٨٠٠ = ١٢١٢ ذوي عيون رمادية او خضر

٢٦٣٢ – (١٠٨٨ + ١٢١٢) = ٣٣٢ ذوي عيون بنية أو سود

مجموعة الرجال ذوي الشعر الأسود:

١٢٢٣ × ٢٨١١ / ٦٨٠٠ = ٥٠٦ ذوي عيون زرق

١٢٢٣ × ٣١٣٢ / ٦٨٠٠ = ٥٦٣ ذوي عيون رمادية أو خضر

١٢٢٣ – (٥٠٦ + ٥٦٣) = ١٥٤ ذوي عيون بنية أو سود

مجموعة الرجال ذوي الشعر الأحمر	
	$٦١٦ \times \frac{٧٨١١}{٦٨٠٠} = ٤٨$ ذوي عيون زرق
	$٦١٦ \times \frac{٣١٣٢}{٦٨٠٠} =$ ٥٣ر٤ ذوي عيون رمادية او خضر
	٦١٦ − (٤٨ + ٥٣ر٤) = ١٤ر٦ ذوي عيون بنية أو سود

بعد ذلك تختبر الفرق بين هذه البيانات المتوقعة (التكرارات النظرية) والبيانات الوقعية الواردة في الجدول السابق .

الصفة	التكرار الواقعي	التكرار النظري	ف	ف٢	ف٢ / التكرار النظري
شعر اصفر وعيون زرق	١٧٦٨	١١٦٩	٥٩٩	٣٥٨٨٠١	٣٠٧
شعر اصفر وعيون خضر	٩٤٦	١٣٠٣	− ٣٥٧	١٢٧٤٤٩	٩٨
شعر اصفر وعيون سود	١١٥	٣٥٧	− ٢٤٢	٥٨٥٦٤	١٦٤
شعر بني وعيون زرق	٨٠٧	١٠٨٨	− ٢٨١	٧٨٩٦١	٧٢
شعر بني وعيون خضر	١٣٨٧	١٢١٢	١٧٥	٣٠٦٢٥	٢٥
شعر بني وعيون سود	٤٣٨	٣٣٢	١٠٦	١١٢٣٦	٣٤
شعر اسود وعيون زرق	١٨٩	٥٠٦	− ٣١٧	١٠٠٤٨٩	١٩٩
شعر اسود وعيون خضر	٧٤٦	٥٦٣	١٨٣	٣٣٤٨٩	٧٠
شعر اسود وعيون سود	٢٨٨	٥٤١	١٣٤	١٧٩٥٦	١١٧
شعر احمر وعيون زرق	٤٧	٤٨	− ١	١	٠٠٠ ×
شعر احمر وعيون خضر	٥٣	٥٣ر٤	− ٤ر	١٦ر	٠٠٠
شعر احمر وعيون سود	١٦	١٤ر٦	٤ر	١ر٩٦	٠٠٠
× الأرقام ضئيلة ويمكن اهمالها					١٠٧٦

وعلى اساس درجات الحرية والتي تساوي في هذا المثال (٤–١) (٣–١) = ٦ او ١٢ تكرار نظري ناقصاً عدد الثوابت التي استخدمت في حساب هذه التكرارات وهي ٦ اي (١٢ – ٦) = ٦ درجات حرية وبالرجوع الى جدول كاي تربيع نجد ان احتمال الحصول على القيمة ١٠٧٦ بطريقة المصادفة احتمال ضعيف جداً ، وبذلك فهذا الاختبار يقيم الدليل على عدم صحة الافتراض بعدم وجود اية علاقة بين لون الشعر ولون العينين وبمعنى آخر بان هذا الاختبار يقيم الدليل على ان تصنيف هذه المجموعة من الذكور تبعاً للون الشعر ليس مستقلا عن تصنيفهم تبعاً للون العينين .

ملاحظات على تطبيق اختبار كاي تربيع

١ – يدلنا اختبار كاي تربيع على احتمال الحصول في حالة المعاينة العشوائية على قيمة لكاي تربيع تساوي او اكثر من القيمة التي حصلنا عليها فعلا بتطبيق الاختبار ، فاذا كان الاحتمال ضعيفاً يكون لنا الحق في الاعتقاد بوجود اختلاف جوهري بين البيانات المشاهدة المتوقعة . ولكنها لا نستطيع ان نسير في الاتجاه العكس فنقول–في حالة ان يكون الاحتمال ليس صغيراً– بصحة الافتراض الذي بدأنا به ، اذا ان كل ما نستطيع قوله ان الاختبار لا يقيم الدليل على عدم صحة الافتراض .

٢ – ان اختبار كاي تربيع ليس مقياساً لدرجة او نوع العلاقة بين الظواهر المختلفة . ان كل ما نستطيع ان نفهمه من هذا الاختبار هو ما اذا كان تصنيف مجموعة من الوحدات تبعاً لكل من الظاهرتين معينتين مستقل عن الاخر او غير مستقل عنه ولكننا لا نستطيع ان نفهم اي شيء فيما يختص بشكل العلاقة بينهما . ولتحديد درجة هذه العلاقة ونوعها لا بد من حساب مقاييس اخرى (مقاييس الارتباط) .

٣ – لا يمكن اجراء الاختبار ما لم تكن التكرارات في شكل مطلق

حيث انه اذا كانت التكرارات في شكل نسبي لا يمكن الحكم على مدى الاختلاف بين البيانات الواقعية والبيانات النظرية . والفرق بين القيمة المطلقة ٣ والقيمة ٤ يختلف كثيرا عن الفرق بين القيمة المطلقة ٣٠٠ والقيمة ٤٠٠ ، بينما اذا اعطينا هذه القيم في شكل نسبي فان الفرق في الحالتين يكون متساوي.

٤ – يجب ان يكون حجم العينة موضوع الدرس ليس صغيراً ، وقد اقترح يول وكندل (Yule and Kendall) الحجم ٥٠ كأقل حجم يمكن دراسته لاجراء الاختبار . وكذلك يجب ان لا يكون عدد الوحدت في كل فئة من الفئات صغيراً وقد جرى العرف على ان خمس وحدات هو اقل عدد يمكن ان يكون في اية فئة من الفئات ويحسن ان يكون عشرة . على ان هذه المشكلة يمكن التغلب عليها بتجمع الوحدات الموجودة في بعض الفئات اذا كان عددها صغيراً خاصة وان الفئات الصغيرة تكون دائماً في اطراف التوزيعات .

٥ – ان من محاسن اختبار كاي تربيع كأداة في الابحاث العلمية ان قيم كاي تربيع المحسوبة لعدة عينات خاصة بنفس الموضوع تحت الدرس ومأخوذة من نفس المجتمع يمكن اضافتها حتى يمكن ان يكون حكمنا افضل عما لو اخذنا كل عينة على حدة .

فاذا افترضنا اننا اخذنا أربع عينات من العمال الصناعيين وصنفنا عمال كل عينة على اساس توظفهم وعلى أساس نوع الصناعة التي يعملون فيها فإننا نحصل على أربع مجموعات في كل عينة – العمال الموظفون والعاملون في الصناعات التي تنتج السلع الرأسمالية ، والعمال الموظفون في الصناعات التي تنتج السلع الاستهلاكية ، والعمال المتعطلون الذين يعملون عادة في الصناعات الرأسمالية ثم المتعطلون والذين يعملون عادة في الصناعات الاستهلاكية .

وعلى أساس البيانات الواقعية التي نحصل عليا وبافتراض عدم وجود أية علاقة بين التوظف ونوع الصناعة سوف نستنتج بيانات نظرية ، ثم نقوم

باجراء اختبار كاى تربيع بالنسبة لكل عينة على حدة ونفترض اننا حصلنا على النتائج التالية :

رقم العينة	عدد درجات الحرية	كاي تربيع
١	١	٣ر٧٥
٢	١	٣ر٦٠
٣	١	٢ر١٢
٤	١	٤ر٢٠
المجموع	٤	١٣ر٦٧

من هذه النتائج ومن جدول كاى تربيع نجد ان العينات الثلاث الأولى تبين عدم وجود فرق جوهري بين البيانات الواقعية والبيانات النظرية باحتمال ٥ ٪ بينما تبين العينة الرابعة وجود فرق جوهري باحتمال ٥ ٪ ولكن ليس باحتمال ١ ٪ ، بينما على أساس مجموع كاى تربيع ومجموع درجات الحرية ٤ تبين وجود فرق جوهري باحتمال ٥ ٪ وكذلك باحتمال ١ ٪ . ولا شك اننا بذلك نستطيع أن نحكم حكما يتفق مع ما نتوقعه حيث انه من المعتقد ان وقع البطالة يكون أكبر في الصناعات الرأسمالية بالمقارنة مع الصناعات الاستهلاكية . وتأكيد وجود فرق جوهري تبعاً لمجموع كاى تربيع يؤيدما نتوقعه ، حيث ان وجود فرق جوهري يقيم الدليل على عدم صحه افتراضنا بعدم وجود علاقة بين التوظف ونوع الصناعة ولإدراك أهمية هـــذه الخاصية لاختبار كاي نورد المثال التالي :

يبين الجدول التالي نتيجة التلقيح ضد الكوليرا في احدى مناطق الدولة .

مثال ٣٦ :

	لم يصب بالمرض	اصيب	المجمـــوع
لقح ضد الكوليرا	٤٣١	٥	٤٣٦
لم يلقح ضد الكوليرا	٢٩٠	٩	٣٠٠
المجمـــوع	٧٢٢	١٤	٧٣٦

نفترض عدم وجود اية علاقة بين التلقيح ضد الكوليرا والاصابة بالمرض أي ان المجموعتين (الذين لقحوا والذين لم يلقحوا) يكونا مجتمعاً واحداً فيه نسبة الاصابة بالمرض $\frac{١٤}{٧٣٦}$ أي $\frac{٧}{٣٦٨}$. وبتطبيق هذه النسبة نحصل على البيانات النظرية التالية :

الذين لقحــوا

$$٤٢٦ \times \frac{٧}{٣٦٨} = ٨{,}٣$$ يصاب بالمرض .

$$٤٣٦ - ٨{,}٣ = ٤٢٧{,}٧$$ لايصاب بالمرض .

الذين لم يلقحوا

$$٣٠٠ \times \frac{٧}{٣٦٨} = ٥{,}٧$$ يصاب بالمرض

$$٣٠٠ - ٥{,}٧ = ٢٩٤{,}٣$$ لايصاب بالمرض

ثم نجري الاختبـــار على الفرق بين البيانات الواقعية والبيانات النظرية كالآتي :

ف۲ التكرار النظري	ف۲	ف	التكرار النظري	التكرار الوقعي	
۰۲٥ر	۸۹ر۱۰	۳ر۳	٤۲۷ر۷	لم يصب ٤۳۱	الذين لقحوا
۳۱۲ر۱	۸۹ر۱۰	۳ر۳	۸ر۳	اصيب ٥	
۰۳۷ر	۸۹ر۱۰	۳ر۳	۲۹٤ر۳	لم يصب ۲۹۱	الذين لم يلقحوا
۹۱۰ر۱	۸۹ر۱۰	۳ر۳	٥ر۷	اصيب ۹	

وعلى أساس درجة حرية واحدة (۲ – ۱) (۲ – ۱) نلاحظ ان قيمة كاى تربيع ۲۸٤ر۳ واقعة بين احتمال ٥ ٪ واحتمال ۱۰ ٪ وبذلك لا نستطيع أن نقرر قراراً فاصلاً في صحة أو عدم صحة افتراضنا .

لذلك يكون من الأفضل اجراء التجربة بالنسبة لعدة عينات من مناطق مختلفة ، فاذا فرضنا اننا حصلنا من هذه العينات النتائج التالية :

رقم العينة	عدد درجات الحرية	كاي تربيع
۱	۱	۲۸ر۳
۲	۱	۳٤ر۹
۳	۱	۰۸ر٦
٤	۱	٥۱ر۲
٥	۱	٦۱ر٥
٦	۱	٥۹ر۱
		٤۱ر۲۸

وعلى أساس ست درجات الحرية يكون احتمال الحصول على هذه القيمة : ٢٨ر٤١ بطريق المصادفة احـــتمال ضعيف جداً (أقل من ١ في الالف) ، وبذلك فهذه القيمة تشكل فرقاً جوهرياً بين البيانات الواقعية والبيانات النظرية ، الأمر الذي يقيم الدليل على عدم صحة الافتراض بعدم وجود علاقة بين التطعيم ضد الكوليرا والاصابة بهذا المرض

اختبار حسن التطابق لتوفيق المنحنى المعتدل :

قمنا في الفصل السابق بتوفيق المنحنى المعتدل (بطريقة المساحات) وتوصلنا بذلك إلى تكرارات تختلف بعض الشيء عن التكرارات الواقعية ، وأشرنا إلى ان الخطوة التالية تكون في اختبار حسن التطابق بين هـــذه التكرارات النظرية والتكرارات المشاهدة الواقعية . والجدول التالي يوضح كيفية اجراء هذا الاختبار :

مثـــال :

التكرار الواقعي	التكرار النظري	ف	ف^{٢}	ف^{٢} / التكرار النظري
٦	٦	صفر	صفر	صفر
١٧	١٢	٥	٢٥	٢ر٠٨
٢٥	٢٦	٩	٨١	٣ر١١
٤٨	٤٨	صفر	صفر	صفر
٦٥	٧٩	١٤	١٩٦	٢ر٤٨
٩٠	١١٣	٢٣	٥٢٩	٤ر٦٨
١٣١	١٤٠,٥	٩,٥	٩٠,٢٥	ر٦٤
١٧٣	١٥١	٢٢	٤٨٤	٣ر٠١
١٥٥	١٤٠,٥	١٤,٥	٢١٠,٢٥	١ر٥٠
١١٧	١١٣	٤	١٦	ر١٤
٧٥	٧٩	٤	١٦	ر٢٠
٥٢	٤٨	٤	١٦	ر٣٤
٢١	٢٦	٥	٢٥	ر٩٥
٩	١٢	٣	٩	ر٧٥
٦	٦	صفر	صفر	صفر
				١٩ر٨٨

وعلى أساس ١٢ درجة حرية (١٥ تكرار نظري – عدد الثوابت التي استخدمت في حساب هذه التكرارات وهي مجموع التكرارات والانحراف

المعياري والوسط الحسابي للتوزيع الواقعي) يتبين من جدول كاي تربيع ان احتمال الحصول على القيمة ١٨٫٩٣ أقل من ٥ ٪ أي احتمال ضعيف ، وبذلك لا يمكن أن نعزو هذا الفرق إلى المصادفة ، فهو اذن فرق جوهري على أساسه يتعين علينا استبعاد الافتراض الذي بدأنا به وهو تماثل التوزيع الخاص باجور العمال موضوع الدرس .

وبالمثل يمكن اختبار حسن التطابق بين التكرارات الواقعية والتكرارات النظرية التي توصلنا اليها بطريقة الاحداثيات الرأسية .

تمارين

١ – يبين الجدول الآتي توزيع الطلبة الناجحين والراسبين :

	صف الاستاذ أ	صف الاستاذ ب	صف الاستاذ ج
الناجحون	٥٠	٤٧	٥٦
الراسبون	٥	١٤	٨

اختبر معنوية الفرق بين نسب النجاح عند الاساتذة الثلاث :

(على أساس مستوى معنوية ٥ ٪ لا يمكن رفض الفرضية بان لا فرق جوهري بين نسب النجاح عند الاساتذة الثلاث) .

٢ – يبين الجدول التالي عـــدد الكتب المعارة في أيام الاسبوع لزائري احدى المكتبات العامة :

الأيام :	الاثنين	الثلاثاء	الاربعاء	الخميس	الجمعة
عدد الكتب :	١٣٥	١٠٨	١٢٠	١١٤	١٤٦

هل هناك فرق جوهري بين عدد الكتب المعارة في هذه الايام ؟

(على اساس مستوى معنوية ٥ ٪ لا يمكن رفض الفرضية بان لا فرق جوهري بين عدد الكتب المعارة في هذه الايام) .

٣ - الآتي توزيع الطلبة تبعاً لدرجاتهم في الرياضيات وفي الطبيعة :

	عالية في الطبيعة	متوسطة في الطبيعة	ضعيفة في الطبيعة
عالية في الرياضيات	٥٦	٧١	١٢
متوسطة » » »	٤٧	١٦٣	٣٨
ضعيفة » » »	١٤	٤٢	٨٥

اختبر الفرضية بان درجات الرياضيات مستقلة عن درجات الطبيعة .

(على اساس مستوى معنوية ٥ ٪ يمكن رفض الفرضية) .

الفصل العاشر

العينات

سبق ان ذكرنا أن الدراسات الاحصائية يمكن أن تجري بالعينة لاستنتاج المقاييس التي نرغب في الوصول اليها والتي تمثل المجتمع الذي سحبت منه العنة. كذلك أشرنا الى أن المقاييس المستنتجة من العينة لا يمكن أن تكون هي نفسها المقاييس الخاصة بالمجتمع اذ لا بد أن يوجد فرق بينهما ، هذا الفرق سميناه بخطأ العينة . فلو أجرينا دراسة عن أطوال الطلبة بالعينة وتبين منها أن متوسط طول الطالب ١٧٢ سم مثلاً فهل معنى ذلك أن متوسط طول الفرد في المجتمع يساوي فعلاً ١٧٢ سم ، وبمعنى آخر هل اذا قمنا بدراسة مجتمع الطلبة بالعد الشامل هل نصل الى هذه النتيجة فعلاً ؟ كذلك اذا قمنا بدراسة نسبة البطالة في المجتمع عن طريق عينة وتبين منها أن هذه النسبة هي ١٥٪ فهل يعني ذلك أن نسبة البطالة في المجتمع هي فعلاً ١٥٪ ؟ .

كذلك بينا فيما سبق أن نتائج العينات يمكن أن تتعرض لأخطاء التحيز ، على أن مثل هذه الأخطاء يجب القضاء عليها بسحب العينة سحباً علمياً سليماً، ولكن بالرغم من ذلك قد لا يتحقق التطابق بين نتائج العينة ونتائج المجتمع الذي سحبت منه لو اجريت الدراسة بالعد الشامل.لذلك يكون من الواجب أن نتعرف على القوانين والعوامل التي تتحكم في الفرق بين المقاييس المستنتجة من العينة ومقاييس المجتمع الفعلية ، أي التي نتحكم في أخطاء

المعاينة . وحتى نستطيع أن نتعرف على هذه القوانين والعوامل يجب أن ندرس التوزيع الاحتمالي لنتائج العينات .

نفترض أن لدينا مجتمعاً معيناً عدد وحداته د واردنا أن نسحب منه العينات الممكنة والتي يكون حجم كل منها = د فما هو عدد هذه العينات ؟ وبمعنى آخر ما عدد الطرق التي يمكن بها اختبار عينة عدد وحداتها = د من مجموع عدد وحدات المجتمع ن ؟ ان عدد العينات التي يمكن سحبها هو ن ق د أي $\frac{ن!}{د!(ن-د)!}$ وللتبسيط نفترض اننا نستطيع أن نسحب ١٠٠٠ عينة من هذا المجتمع على أساس هذا الحجم للعينة . نقوم بعد ذلك بحساب المتوسط لكل عينة وبذلك يتكون لدينا ١٠٠٠ وسط يمثل جميع العينات التي أمكن سحبها من هذا المجتمع . نوزع هذه المتوسطات في توزيع تكراري ، وبذلك نحصل على عدد العينات التي يكون متوسطها واقعاً في كل فئة من فئات التوزيع . هذا التوزيع نسميه بتوزيع المعاينة ويمكن تعريفه بانه توزيع متوسطات جميع العينات التي يمكن سحبها من مجتمع معين على أساس حجم معين لكل عينة من هذه العينات . ويتميز هذا التوزيع بالصفات الآتية :

١ — انه يكون توزيعاً معتدلاً أو قريباً جداً من الاعتدال .

٢ — ان الوسط الحسابي لهذا التوزيع يكون هو نفسه المتوسط الحسابي للمجتمع الذي سحبت منه العينات أو قريباً جداً منه . وبذلك نستطيع أن نقول ان متوسط متوسطات جميع العينات التي يمكن سحبها من مجتمع معين يكون هو نفسه المتوسط الواقعي للمجتمع الذي نسحب منه هذه العينات .

٣ — حيث ان توزيع المعاينة يكون توزيعاً معتدلاً بذلك نستطيع أن نطبق خاصية هذا التوزيع فنقول ان ٦٨ ٪ من العينات التي يمكن سحبها من

مجتمع معين لا يمكن أن يزيد متوسطها أو يقل عن متوسط المجتمع الاصلي إلا بمقدار ١ ع ، وان ٩٥٫٤ ٪ من العينات التي يمكن سحبها من مجتمع معين لا يمكن أن يزيد متوسطها أو يقل عن الوسط الواقعي للمجتمع إلا بمقدار ٢ ع ، وان ٩٩٫٧ ٪ من العينات التي يمكن سحبها من مجتمع معين لا يمكن أن يزيد متوسطها أو يقل عن الوسط الواقعي إلا بمقدار ٣ ع .

وبشكل آخر نستطيع أن نقول ان احتمال أن يكون المتوسط المقدر من عينة أكبر أو أقل من المتوسط الواقعي للمجتمع بمقدار ١ ع هو ٦٨ ٪ . وان احتمال أن يكون المتوسط المقدر من عينة أكبر أو اقل من المتوسط الواقعي للمجتمع بمقدار ٢ ع هو ٩٥٫٤ ٪ . وان احتمال ان يكون المتوسط المقدر من عينة أكبر أو أقل من المتوسط الواقعي للمجتمع بمقدار ٣ ع هو ٩٩٫٧ ٪

٤ - ان الانحراف المعياري الذي تكلمنا عنه في الفقرة السابقة هو الانحراف المعياري لتوزيع المعاينة ، ونلاحظ ان هذا التوزيع هو توزيع نظري فقط أي لا يوجد لدينا عملياً حيث انه يمثل متوسطات جميع العينات التي يمكن سحبها من المجتمع ونحن في الواقع لا نسحب إلا عينة واحدة من هذا المجتمع . وبناء على الصفة السابقة لتوزيع المعاينة نستطيع أن نقول ان هذه العينة التي سحبناها تقع ضمن هذا التوزيع النظري ونستطيع أن نحدد مكان وقوعها فيه باحتمالات مختلفة ، فباحتمال ٦٨ ٪ تقع هذه العينة في القطاع الاول (س‾ + ١ ع ، س‾ - ١ ع) وباحتمال ٩٥٫٤ ٪ تقع في القطاع الثاني (س‾ + ٢ ع ، س‾ - ٢ ع) وباحتمال ٩٩٫٧ ٪ تقع في القطاع الثالث (س‾ + ٣ ع ، س‾ - ٣ ع) .

ان الانحراف المعياري لتوزيع المعاينة يتوقف اولا على الانحراف المعياري للمجتمع الاصلي الذي نسحب منه العينات وعلى حجم العينات التي نسحبها والرسم التالي يوضح لنا كيف ان الانحراف المعياري لتوزيع المعاينة يقل

تدريجياً كلما زاد عدد الوحدات التي نسحبها في كل عينة . أي ان الانحراف المعياري لتوزيع المعاينة يتناسب طردياً مع الانحراف المعياري للمجتمع وعكسياً مع حجم العينات التي تسحب منه . وبالبحث الرياضي وجد ان الانحراف

المعياري لتوزيع المعاينة = $\frac{ع}{\sqrt{ن}}$ ، على أن القانون الاصلي

للانحراف المعياري لتوزيع المعاينة هو = $\frac{ع}{\sqrt{ن}} \sqrt{\frac{ن - د}{ن - ١}}$

ولكن عندما يكون حجم المجتمع لانهائياً تقترب قيمته $\sqrt{\frac{ن - د}{ن - ١}}$

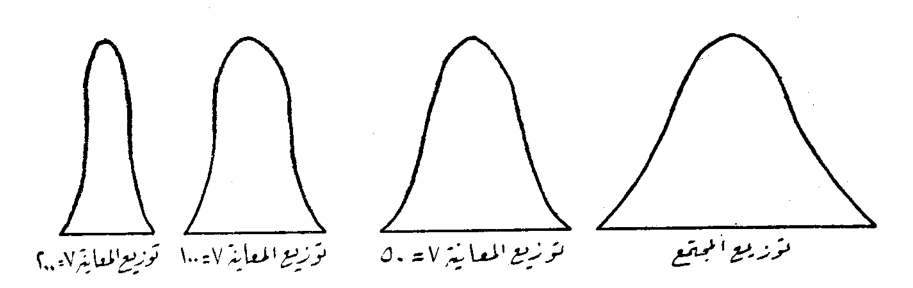
من الواحد الصحيح وبذلك يكون الانحراف المعياري لتوزيع المعاينة =

$\frac{ع}{\sqrt{ن}}$ وهو القانون الشائع الاستعمال في الدراسات الاحصائية . ويسمى

الانحراف المعياري لتوزيع المعاينة تمييزاً له عن الانحراف المعياري للمجتمع بالخطأ المعياري . ويقيس لنا هذا المقياس مدى التشتت بين متوسطات العينات نتيجة عوامل الصدفة في سحب هذه العينات ، اي انه بالتالي يقيس الاخطاء العشوائية المتوقعة التي يتعرض له المتوسط المقدر من عينة . وفي الحالات التي لا يتوفر فيها الانحراف المعياري للمجتمع الذي سحبت منه العينة يمكن ان نستعيض بدلا عنه بالانحراف المعياري لوحدات العينة نفسها ، ومن الواضح ان هذا الانحراف المعيارى لا يمكن ان يكون هو نفسه الانحراف المعياري للمجتمع ، الا أن الفرق بينهما يكون ضئيلا إذا أمكن سحب العينة بحيث تمثل الاختلافات بين وحدات المجتمع تمثيلاً صحيحاً .

وخلاصة ما تقدم نقول انه عندما نحسب الخطأ المعياري للعينة يكون

ذلك باحتمال معين أي بدرجة ثقـة ليست كاملة ، كما أن هذا الخطأ يكون احتمال الزيادة عن الوسط الواقعي للمجتمع واحتمال النقص عنه . وقد اختلف



الاحصائيون في تحديد حدود الثقة ، الا اننا سنعتبر ان وقوع وسط العينة (أو قيمة وحدة معينة في حالة التوزيع المعتدل لمجتمع احصائي) في حدود درجة الثقة ٩٥٪ يجعلنا نستنتج انه ليس هناك ما يدعو الى الشك في عشوائية العينة أي أن الفرق بين متوسطها ومتوسط المجتمع هو فرق راجع الى الصدفة أي أنه غير معنوي أو غير جوهري . واذا وقع متوسط العينة بين درجتي الثقة ٩٥٪ و ٩٩٪ يكون هناك شك في عشوائية العينة . واذا وقع متوسط العينة خارج درجة الثقة ٩٩٪ نؤكد عدم عشوائية العينة أي أن الفرق بينها وبين المجتمع فرق جوهري أي معنوي أي لا يمكن أن يكون راجعاً لمجرد الصدفة. وعلى أساس درجة الثقة ٩٩٪ يكون حدي الثقة هما $\bar{س}$ + ٢٫٥٨ع، $\bar{س}$ − ٢٫٥٨ ع . أما اذا اعتبرنا درجة الثقة ٩٥٪ يكون حدي الثقة هما :

$\bar{س}$ − ١٫٩٦ ع ، $\bar{س}$ − ١٫٩٦ ع

استخدام متوسط العينة في تقدير متوسط المجتمع :

مثال ٣٨ :

نفترض اننا أخذنا عينة من الأطفال مكونة من ١٠٠ طفل تبين منها ان

متوسط الوزن هو ٣٠٫٥ كيلو بانحراف معياري = ٩ كيلو ونريد أن نجد حدود متوسط وزن الاطفال من نفس العمر في المجتمع .

$$\text{الخطأ المعياري للعينة} = \frac{٩}{\sqrt{١٠٠}} = ٫٩$$

وبذلك بدرجة ثقة ٩٥٪ يكون متوسط وزن الأطفال في المجتمع واقعاً في فترة الثقة :

$٣٠٫٥ \pm ١٫٩٦ \times ٫٩$

أي بين $٣٠٫٥ \pm ١٫٧٦٤$

أي بين ٣٢٫٢٦٤ و ٢٨٫٧٣٦ كيلو

وبدرجة ثقة ٩٩٪ يكون متوسط وزن الأطفال في المجتمع واقعاً في فترة الثقة :

$٣٠٫٥ \pm ٢٫٥٨ \times ٫٩$

أي بين $٣٠٫٥ \pm ٢٫٣٢٢$

أي بين ٣٢٫٨٢٢ و ٢٨٫١٧٨ كيلو

ونلاحظ أن المدى في الحالة الثانية أوسع منـــه في الحاله الاولى وذلك طبيعي حيث أننا في الحالة الثانية نكون أكثر وثوقاً من النتيجة الـتي نصل اليها بالمقارنة مع نتيجة الحالة الأولى .

اختبار عشوائية العينة أي اختبار معنوية الفرق بين متوسط العينـة ومتوسط المجتمع :

اذا توفر لنا متوسط المجتمع الذي سحبت منه عينة ما فاننا نستطيع بمقارنة هذا المتوسط بالمتوسط المقدر من العينة على أساس الدرجات المعيارية أن نختبر عشوائية العينة أي مدى تمثيلها للمجتمع وبمعنى آخر نستطيع أن نحكم على معنوية الفرق بينها وبين المجتمع ، ويكون لذلك أهمية كبيرة حيث اذا تبينا عدم وجود فرق معنوي بين العينة والمجتمع الذي سحبت منه فاننا نطمئن اليها ونستمر في دراستها .

مثال ٣٩ :

اذا كان متوسط الأجر في صناعة ما هو ٩٦ ليرة في الاسبوع واخذت عينة من المشتغلين في هذه الصناعة وكان توزيعهم حسب فئات الأجر كالآتي :

فئات الاجر	عدد العمال	م	ح	ح ك	ح ٢ ك
٧٠ —	٧	٧٥	— ٢	— ٢٤	٢٨
٨٠ —	١٣	٨٥	— ١	— ١٢	١٣
٩٠ —	٢٥	٩٥	صفر	صفر	
١٠٠ —	١٨	١٠٥	+ ١	+ ١٨	١٨
١١٠ —	١٢	١١٥	+ ٢	+ ٢٤	١٨
١٢٠—١٣٠	٦	١٢٥	٣٩	+ ١٨	٥٤
المجمـــوع	٨١	—	—	٣٣	١٦١

$$\bar{س} = ٩٥ + \frac{٣٣}{٨١} \times ١٠ = ٩٩٫٠٧ \text{ ليرة}$$

$$ع = ١٠ \sqrt{\frac{١٦١}{٨١} - \left(\frac{٣٣}{٨١}\right)^{٢}} = ١٣٫٤٩ \text{ ليرة}$$

∴ الانحراف المعياري لتوزيع المعاينة (الخطأ المعياري) =

$$\frac{١٣٫٤٩}{\sqrt{٨١}} = ١٫٤٩٨$$

∴ الدرجة المعيارية = $\frac{٩٩٫٠٧ - ٩٦}{١٫٤٩٨}$ = ٢٫٠٥ .

∴ متوسط العينة واقع بين الدرجة المعيارية ١٫٩٦ والدرجة المعيارية ٢٫٥٨ . وبذلك يكون هناك شك في عشوائية العينة .

ويمكن أن نصل الى نفس النتيجة بطريقة اخرى إذ نستطيع أن نحسب فترة الثقة بدرجة ٩٥ ٪ :

$$٩٩٫٠٧ \pm ١٫٩٦ \times ١٫٤٩٨$$

اي بين $٩٩٫٠٧ \pm ٢٫٩٣٦١$

اي بين ١٠٢٫٠٠٦١ ، ٩٦٫١٣٣٩ .

وحيث أن المتوسط الواقعي لا يقع في هذا المدى فـلا يمكن التأكد من عشوائية العينة . وبدرجة تقة ٩٩٪ يكون المدى :

$$٩٩٫٠٧ \pm ٢٫٥٨ \times ١٫٤٩٨$$

أي بين $٩٩٫٠٧ \pm ٣٨٫٦٤٨$

أي بين ١٠٢٫٩٣٤٨ ، ٩٥٫٢٠٥٢

وحيث أن المتوسط الواقعي (٩٦) يقع في هذا المدى فيكون بذلك محصوراً بين درجة الثقة ٩٥٪ ودرجة الثقة ٩٩٪ ، أي المدى الذي يجعلنا نشك في عشوانيه العينة .

وبنفس الطريقة السابقة يمكن اختيار معنوية الفرق بين متوسط مجتمع معين ومتوسط جزء من هذا المجتمع على افتراض أن هذا الجزء يمثل عينة عشوائية من المجتمع لا تختلف في صفاتها عن صفات المجتمع كله . فاذا تبين لنا عدم وجود فرق معنوي نستطيع أن نصل الى النتيجة بان هذا الجزء يمكن ان يعتبر عينة ممثلة تمثيلاً صادقاً للمجتمع،أما اذا تبين لنا وجود فرق جوهري نستطيع أن نصل الى النتيجة بان هناك عوامل معينة تجعل هذا الجزء يختلف اختلافاً جوهرياً عن المجتمع وبالتالي لا يمكن اعتباره عينة عشوائية سليمة من هذا المجتمع ، ويمكننا أن نستفيد من ذلك في مقارنه صفات المجتمعات الاحصائية التي ندرسها بصفات أجزائها، فمثلاً نستطيع أن نقارن بين متوسط انتاج مصنع معين وانتاج جميع المصانع الشبيهة ، وبين انتاج منطقة معينة في الدولة من محصول معين وانتاج الدولة باكملها من هذا المحصول ، وبين متوسط وزن الاطفال في مدرسة معينة وبين متوسط وزن الأطفال جميعاً في نفس العمر ... الخ

مثال ٤٠ :

كان متوسط محصول الفدان الواحد من القطن في احدى المحافظات عام ١٩٦٢ = ٤٢٠ رطلا وكان الانحراف المعياري لغلة الفدان ٩٤ رطلا ، فاذا علمنا أن المتوسط العام لغلة الفدان في الدولة هو ٤٣١ رطلا ، اختبر معنوية الفرق بين المتوسطين اذا كانت المساحة المزروعة قطناً في هذه المحافظة هي ٢٠٠ فدان .

$$\text{الخطأ المعياري} = \frac{٩٤}{\sqrt{٢٠٠}} = ٦٫٧ \text{ رطل}.$$

$$\text{الدرجة المعيارية} = \frac{٤٣١ - ٤٢٠}{٦٫٧} = \frac{١١}{٦٫٧} = ١٫٦ \text{ رطل}.$$

وبذلك يكون متوسط المحافظة داخل درجة ثقة ٩٥ ٪ ، وبذلك لا يكون هناك ما يدعو إلى رفض الفرض القائل بان أرض هذه المحافظة المزروعة قطناً تعتبر عينة عشوائية من أراضي الدولة المزروعة بهذا المحصول .

وبطريقة اخرى يكون اقصى فرق بين المتوسطين بدرجة ثقة ٩٥ ٪ :

= ٦٫٧ × ١٫٩٦ = ١٣٫١٣٢ رطل .

وحيث ان الفرق بين المحافظة والمجتمع هو ١١ رطل فقط أي انه أقل من ١٫٩٦ خطأ معياري ، بذلك يكون الفرق في حدود ما تأتي به الصدف العشوائية .

ويفيدنا الاختبار السابق في الاجابة على السؤال – هل لعامل معين أثر على جزء من المجتمع – أي هل له تأثيره في اختلاف متوسط الجزء عن متوسط المجتمع؟ ولا شك انه لا يجب ان نتسرع ونحكم بذلك فقد يكون الفرق راجعاً لمجرد الصدفة العشوائية . ولذلك على اساس هذا الاختبار نستطيع أن نحكم في الموضوع . فاذا لم يكن هناك أي شك في عشوائية العينة أي إذا كان متوسطها لا يختلف عن الوسط العام للمجتمع زيادة او نقصاً إلا بمقدار ١٫٩٦ خطأ معياري في الحد الأقصى فلا يكون لدينا أي شيء يدعو إلى الاعتقاد بوجود تأثير لهذا العامل . أما إذا كان هناك شك في عشوائية العينة ، أي إذا كان متوسطها بين درجة ثقة ٩٥ ٪ و ٩٩ ٪ يكون هناك ترجيح لوجود

تأثير لهذا العامل . وإذا كان في حكم المؤكد أن العينة غير عشوائية أي إذا خرج متوسطها عن درجة الثقة ٩٩ ٪ يكون في حكم المؤكد وجود تأثير لهذا العامل ، على انه يجب أن نلاحظ عدم وجود أي عامل آخر يمكن أن يكون له تأثير أي لا بد أن تتشابه الظروف بين الجزء والمجتمع ولا يكون هناك من فرق بينهما إلا بالنسبة لهذا العامل فقط وهو الذي ندرس تأثيره .

مثـــال ٤١ :

إذا علمنا ان متوسط انتاج العامل الواحد في صناعة ما هو ٦٠ قطعة بانحراف معياري ٧ قطع – وإذا علمنا أن عينة من ٩٠ عامل دربوا على العمل بطريقة أفضل فأصبح متوسط انتاجهم ٦٥ قطعة ، فهل نستنتج من ذلك أن هذا التدريب كان له أثره على تحسين الانتاج :

بدرجـة ثقة ٩٥ ٪ يكون المدى بين ٦٠ $\pm \frac{٧}{\sqrt{٩٠}} \times ١٫٩٦$

أي بين ٥٨٫٥٥ ، ٦١٫٤٥

وبدرجة ثقة ٩٩ ٪ يكون المدى بين ٦٠ $\pm \frac{٧}{\sqrt{٩٠}} \times ٢٫٥٨$

أي بين ٥٨٫١ ، ٦١٫٩

وحيث ان متوسط الانتاج بعد التدريب أصبح ٦٥ قطعة ، إذا نستطيع أن نستنتج أن التدريب يؤدي فعلاً إلى تحسين الانتاج .

اختبار الفرق بين متوسطي عينتين :

قد يتوفر لدينا بيانات عن مجموعتين متماثلتين في جميع الظروف فيما عدا

عامل واحد يمكن أن يكون قد اثر في احدى المجمرعتين دون الاخرى ، او قد أثر فيهما بدرجات متفاوتة ، ونريـــد أن نتثبت فعلاً من تأثير هذا العامل ، وفي هذه الحالة لا يجب أن نتسرع ونقول أن الفرق بين المجموعتين أو العينين راجع فعلا إلى هذا العامل ، إذ يجب علينا أولا أن نتحقق من امكان أن يكون الفرق بينهما راجع إلى مجرد الصدفة ، فاذا استطعنا أن نستبعد هذا الفرض (فرض الصدفة) أمكننا أن نؤكد تأثير هذ العامل .

لنفترض مثلا ان إحدى المؤسسات قـــامت بتطبيق طريقتين من طرق الانتاج وتبين من اتباعهما أن متوسط جودة الانتاج قد اختلف ، فهل يكون هذا الاختلاف راجع إلى الصدفة أو هو اختلاف جوهري يجعلنا نؤكد تفضيل إحدى الطريقتين على الاخرى ؟ كذلك إذا أخذنا عينتين من الأراضي الزراعية واستعملنا في إحداهما نوعاً من السماد وفي الاخرى نوعاً آخر أو تركناها بدون تسميد وفي نهاية الموسم لاحظنا ان متوسط انتاج العينة الاولى اختلف عن متوسط انتاج العينة الثانية فهل يكون الفرق راجعاً إلى الصدفة أو هو فرق جوهري ، الأمر الذي يجعلنا نؤكد تفضيل نوع على آخر او تأثير السماد على الانتاج إذا لم تكن الأرض في العينة الاخرى قد سمدت .

والسؤال الذي نريد الاجابة عليه في مثل هذه الحالة ، هو هــــل يمكن اعتبار المجموعتين او العينتين منتميتان إلى مجتمعين لهما نفس الوسط الحسابي او لا ؟ وبمعنى آخر ان اختلاف هذا العامل المؤثر لا يقضي على احتمال كون العينتين من مجتمعين لهما نفس الوسط الحسابي ، وبذلك يكون الاختلاف بينهما راجعاً إلى الصدفة ، او انهما ينتميان إلى مجتمعين مختلفين تماماً ، وبذلك نؤكد تأثير هذا العامل .

إذا رمزنا إلى متوسط العينة الاولى بالرمز $\bar{س}_1$ ، ومتوسط العينة الثانية بالرمز $\bar{س}_2$ والفرق بينهما بالرمز ف ، يكون هذا الفرق هو إحدى الفروق التي يتكون منها توزيع لجميـــع الفروق المحتملة بين العينات التي يمكن اخذها

من المجتمعين الأصليين والتي يكون حجم كل عينة من المجتمع الاول $ن_1$ ومن المجتمع الثاني $ن_2$. وفي حالة تساوي الوسطين الحسابيين للمجتمعين فان الوسط الحسابي للفروق المحتملة بينهما لا بد ان يساوي صفراً حيث ان كل فرق موجب لا بد ان يقابله فرق آخر سالب . ويكون توزيع هذه الفروق توزيعاً معتدلاً وسطه الحسابي يساوي صفراً وتباينه يساوي مجموع تباين الوسطين الحسابيين $\bar{س}_1$ ، $\bar{س}_2$.

وحيث ان تباين المجتمع الأول للعينات = $\frac{ع_1^2}{ن_1}$ وتباين المجتمع الثاني للعينات = $\frac{ع_2^2}{ن_2}$ يكون تباين توزيع الفروق بينهما = $\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}$

أي ان $ف^2 = \frac{ع_2^2}{ن_1} + \frac{ع_2^2}{ن_2}$

ويكون بذلك الخطأ المعياري لتوزيع الفروي = $\sqrt{\frac{ع_1^2}{ن} + \frac{ع_2^2}{ن_1}}$

وحيث ان توزيع مجتمع الفروق توزيع معتدل وسطه الحسابي = صفر ، نستطيع ن نطبق قواعد التوزيع المعتدل فنقول ان :

بدرجة ثقة ٩٥ ٪ تكون فترة الثقة = صفر $\pm$ ١٫٩٦

$\sqrt{\frac{ع^2}{ن_1} + \frac{ع_2^2}{ن_2}}$

وبدرجة ثقة ٩٩ ٪ تكون فترة الثقة = صفر $\pm$ ٢٫٥٨

$\sqrt{\frac{ع_1^2}{ن_1} + \frac{ع_2^2}{ن_2}}$

وبذلك إذا وقع الفرق بين العينين داخل فترة الثقة الاولى لا يكون هناك ما يبعث على الشك في ان تكون العينتان مأخوذتان من مجتمعين لهما نفس الوسط الحسابي أي ان الفرق بينهما غير جوهري ويمكن أن يكون راجعاً إلى الصدفة . أما إذا وقع الفرق بين العينتين بين فترة الثقة الاولى وفترة الثقة الثانية يكون هناك احتمال ان تكون العينتان مأخوذتان من مجتمعين لهما نفس الوسط الحسابي . أمـــا إذا وقع الفرق بين العينتين خارج فترة الثقة الثانية نستطيع أن نؤكد ان العينتين غير مأخوذتين من مجتمعين لهما نفس الوسط الحسابي ، وبذلك يكون الفرق بينهما فرق جوهري لا يمكن أن يكون مصدره الصدفة أي اننا نستطيع أن نؤكد تأثير العامل الذي نقوم بدراسته .

مثـــال ٤١ :

الجدول الآتي يبين توزيـع المتعطلين في منطقتين حسب أعمـــارهم ، والمطلوب بحث ما إذا كان هناك اختلاف جوهري بين متوسط عمر المتعطلين في المنطقتين .

العمر بالسنوات الـــكاملة	١١ –	٢٠ –	٣٠ –	٤٠ –	٥٠ – ٦٠	المجموع
عدد المتعطلين في المنطقة أ	٤	١٧	١٨	١١	٨	٥٨
عدد المتعطلين في المنطقة ب	١٠	٢١	٢١	٢٧	١٠	٨٩

نحسب أولا متوسط عمر المتعطل في المنطقتين :

$\bar{س}_{١}$ للمنطقة الاولى = ٣٥ر٣٤ سنة

$\bar{س}_{٢}$ للمنطقة الثانية = ٣٥ر٦٧ سنة

الانحراف المعياري للعمر في المنطقة الاولى ١١ر٤٤ سنة

الانحراف المعياري للعمر في المنطة الثانية = ١١ر٩٧ سنة

$$\therefore \text{ع}_\text{ف} = \sqrt{\frac{{١١٫٤٤}^2}{٥٨} + \frac{{١١٫٩٧}^2}{٨٩}}$$

$$= ١٫٩٦ \text{ سنة .}$$

$\therefore$ فترة الثقة الاولى = $\pm$ ١٫٩٦ × ١٫٩٦ = ٣٫٨٤١٦

وحيث ان الفرق بين متوسطي العينين = ٣٥٫٦٧ − ٣٥٫٣٤ = ٠٫٣٣ سنة وهو واقع ضمن فترة الثقة الاولى .

$\therefore$ الفرق بينهما من المؤكد أن يكون راجعاً إلى الصدفة .

مثال ٤٢ :

من مجموع العمال المشتغلين في احـــدى الصناعات اخذت عينة من العمال الذكور عددهم ٢٠٠ عامل ووجد ان متوسط اجرهم = ٣٥ ليرة بانحراف معياري ٦ ليرات ، وعينة اخرى من الاناث عددهن ٥٠ عاملة ووجد أن متوسط اجورهن ٣٠ ليرة بانحراف معياري ٣ ليرات والمطلوب اختبار الفرق بين متوسطي اجر العينين .

$$\text{ع}_\text{ف} = \sqrt{\frac{٢٥}{٢٠٠} + \frac{٩}{٥٠}}$$

$\therefore$ فترة الثقة الاولى = صفر $\pm$ ١٫٩٦ × ٫٥٥ = $\pm$ ١٫٠٨

وفترة الثقة الثانية = صفر $\pm$ ٢٫٥٨ × ٫٥٥ = $\pm$ ١٫٤٢

وحيث ان الفرك بين متوسطي الأجر = ٣٥ − ٣٠ = ٥ ليرات ، وهو يقع خارج فترة الثقة الثانية يكون في حكم المؤكد اختلاف اجور العمال الذكور والعمال الأناث اختلافاً جوهرياً .

تحليل نتائج العينات في حالة الظواهر الوصفية :

مبادىء اولية في نظرية الاحتمالات :

عندما نتكلم عن نسبة ظاهرة ما ، مثلا نسبة الامية او نسبة البطالة او نسبة المصابين بمرض معين .. الخ نعنى في الواقع احتمال حدوث اي ظاهرة من هذه الظواهر . فاذا قلنا ان نسبة الامية هي ٢٥ ٪ فهل نعني بذلك ان كل ١٠٠ شخص لا بد ان يوجد بينهم ٢٥ أميين ، الواقع اننا نعني انه من كل ١٠٠ شخص يحتمل ان نجد بينهم ٢٥ اميين ، واذا عممنا هذا على المجموع يكون متوسط عدد الاميين من كل ١٠٠ شخص هو ٢٥ . وبذلك تكون النسبة ٢٥ ٪ هي متوسط لتعميم احتمال معين على مجموع معين من الوحدات . وحتى نفهم ذلك يجب ان ندرك الفرق بين المعنى الرياضي والمعنى الاحصائي للاحتمالات .

فالشخص العادي عندما يتكلم على احتمال ظاهرة ما إنما يعني عدم تأكده من حدوثها أو عدم حدوثها وهو بذلك يدلي برأيه في شكل مبهم غير واضح حيث أن رأيه هذا لا يفرق بين الاحتمال الضعيف أو القوي ، وبمعنى آخر لا يقوم هذا الرأي على أي أساس من القياس الدقيق .

على ان الشخص العادي قد يحاول اعطاء رأيه بعض الأهمية فيقول مثلاً ان احتمال وقوع حدث معين هو ٩٩ ٪ إظهاراً لدرجة ثقته في وقوع الحدث او ٥ ٪ إظهاراً لدرجة عدم ثقته في وقوعه ، إلا ان ذلك لا يعني ان هذا الشخص قد استعمل مقياسا دقيقاً للتعبير عن احتمال هذا الحدث فهذا القياس ليس الا مجرد تقدير شخصي لا أساس رياضي له .

اما الاحتمال في المعنى الرياضي فيتميز عن الاحتمال بمعناه العادي السابق في انه يمكن قياسه وتحديده بدقة إذا توفرت لدينا معلومات عن الحدث المراد تقدير احتمال وقوعه وان هذا القياس ليس قائماً على أساس شخصي وإنما هو قياس موضوعي بمعنى انه لا يمكن ان يختلف اثنان في قياس احتمال وقوع حدث معين أو عدم وقوعه إذا أقاما هذا القياس على أساس نفس المعلومات اللازم توفرها .

ويقاس الاحتمال بمقياس يبدأ بالصفر في ناحيه وبالواحد الصحيح في ناحية أخرى . ويدل الواحد الصحيح في قمة المقيـاس على التأكد التام من الحدث بحيث لا يكون هناك أي شك يساورنا عندما نتكلم عن هذا الحدث ، ومثلاً على ذلك احتمال أن يموت الانسان يوماً هو احتمال كامـــل يساوي الواحد الواحد الصحيح حيث لا مجال للشك في وقوع هــذا الحدث ، وفي هذه الحالة يعبر الرياضيون عن هذا الاحتمال كالآتي : (ح = ١) أما اسفل المقياس والمرقم بالصفر فيعـبر عن الاستحالة التامة للحدث وفي هذه الحالة لا يكون هناك أي شك بل تأكد تام من عدم وقوع الحدث ، ومثلاً على ذلك احتمال أن يستطيع انسان ما ان يعبر سباحة المحيط الاطلسي = صفر حيث ان الفشل في المحاولة أمر مؤكد وبذلك تكون ح = صفر .

والواقع انه لو كانت كل الاحداث التي يجابهها الانسان في حياته من الممكن البت فيها بشكلٍ قاطع ولا مجال للشك فيها بأي حال من الاحوال لما كان هنـــاك دافع الى دراسة موضوع الاحتمالات إلا ان الحياة تواجه الانسان بمشاكل لا حصر لها والتي لا يستطيع ان يدلي فيها برأيه بشكل قاطع بحيث تكون ح = ١ أو ح = صفر . فالطبيب مثلاً يعلم ان البنسيلين يفيد مرض معين عند مريض ما ولكنه لا يستطيع أن يؤكد ان استعماله سوف يشفي هذا المريض وبذلك لو كان الطبيب دقيقاً لما استطاع أن يقول ان احتمال شفاء المريض باستعمال البنسيلين = ١ أي انه يثق ذلك ١٠٠ ٪ .

والمشكلة التي تواجهنا الآن هي كيف نستطيـــع ان نقيس احتمال حدث معين اذا لم يكن من النوعين السابق الاشارة اليهما في الفقرات السابقة .

إذا رمينا قطعة من النقود الى أعلى وكانت كاملة التوازن وأجري رميها بدون أي تحيز فما هو احتمال أن تسقط على أحد وجهيها عندما تستقر ؟ من الواضح انه لا يمكن أن يحدث إلا أحد أمرين حيث انها اما أن تسقط على الوجه وإما أن تسقط على الظهر وبذلك يكون الاحتمالين متعادلين ، وعلى أساس جعل مقياس التأكد = ١ صحيح فان احتمال سقوط قطعة النقود على

أحـد وجهيها $= \frac{١}{٢}$ حيث اننا قسمنا الواحد بالتساوي بين الاحتمالين المتعادلين

وبمعنى آخر ان احتمال سقوط قطعة النقود على الوجه $= \frac{١}{٢}$ واحتمال سقوطها

على الظهر $\frac{١}{٢}$ ومجموع الاحتمالين = ١ ويدل هذا المجموع على تأكدنا التام من سقوطها حتما على أحد الوجهين وهذا المجموع قسمناه بالتساوي بين الاحتمالين الممكن وقوعهما. على انه لا يجب أن نفهم من هذا اننا اذا رمينا قطعة النقود مرتين فان الوجه سوف يظهر في إحداهما وسوف يظهر الظهر في المرة الاخرى فقد نرمي القطعة عشر مرات دون أن يظهر أحد الوجهين بتاتاً . وبذلك يكون

المعنى الدقيق لقولنا ان احتمال سقوط قطعة النقود على الوجه يساوي $\frac{١}{٢}$

هو اننا إذا رمينا قطعة النقود عدداً كبيراً جداً من المرات فان الوجه سوف يظهر في نصف هذه المرات (تقريباً) وسنعود الى شرح ذلـــك بالتفصيل فيما بعد .

وعلى نفس الأساس إذا رمينا زهرة نرد كاملة التوازن ، ودون أي تحيز

فان احتمال ظهور أي وجه لها $= \frac{١}{٦}$ حيث اننا متأكدين من اظهارها لوجه

واحد من الأوجه الستة وبذلك يكون مجموع الاحتمالات التي يمكن أن تظهرها زهرة النرد = ١ ، وهذا المجوع يقسم بالتساوي بين الست أوجه التي تكون الزهرة (حيث اننا بالطبع ستبعد أن تقع الزهرة على أي ركن من أركانها أو اي حرف من أحرفها) .

كذلك إذا خلطنا مجموعة من ورق اللعب بحيث يكون لكل ورقة نفس الفرصة في الظهور عند السحب منه ، وإذا لم يكن الساحب متحيزاً فان

احتمال ظهور أي ورقة $= \frac{١}{٥٢}$ حيث اننا متأكدين من ظهور ورقة واحدة من

٥٢ ورقة ولما كان مجموع الاحتمالات التي يمكن أن تظهر = ١ وهذا المجموع مقسم بالتساوي (نظرا لعدم التحيز) بين الـ ٥٢ ورقة فان احتمال ظهور ورقة واحدة = $\frac{١}{٥٢}$. كما ان احتمال ظهور اية صورة = $\frac{١٢}{٥٢} = \frac{٣}{١٣}$ حيث انه من الـ ٥٢ احتمال هناك ١٢ احتمال كل منهما يمكن أن يكون صـورة .

ويمكننا أن نعبر بشكل آخر عما تقدم فنقول اننا اذا رمينا زهرة نرد ٦٠٠ مرة فاننا نتوقع ظهور وجه معين ١٠٠ مرة منها ، وتوقعنا لا يعـني تأكدنا من ذلك تماماً بل بالعكس فإن تعجبنا يكون شديدا لو اتفقت نتيجة التجربة مع ما نتوقعه اتفاقاً تاماً . وبذلك يمكننا أن نقيس الاحتمال بطريقة اخرى وذلك بعدد المرات التي يحدث فيها الحدث من مجمـوع المحاولات التي أجريناها ، على أن يكون عدد المحاولات كبيراً حتى لا يكون للصدفة أي أثر على النتيجة التي نتوصل اليها وعلى ذلك يمكن وضع القاعدة التالية :

$$\text{احتمال حدث ما} = \frac{\text{عدد مرات وقوع الحدث}}{\text{عدد المحاولات}}$$

ونضرب لذلك مثلاً : إذا قام جراح باجراء عملية جراحية معينة على ٢٠٠ شخص وتوفي ١٦ منهم نتيجة العملية نستطيع أن نقول ان احتمال الوفاة بسبب اجراء هذه العملية = $\frac{١٦}{٢٠٠}$ = ٠٫٠٨

ونظرية الاحتمالات مع تعقيدها تنبني على قاعدتين أساسيتين وهما جمع الاحتمالات وضربها .

جمع الاحتمالات :

أذا كان احتمال وقوع حدث معين = $ح_١$ واحتمال وقوع حدث آخر =

$ح_٢$ وكان حدوث أيهما مانعاً من حدوث الآخر ، فعلى أساس وقوع أي من الحدثين يكون بمثابة نجاح فان احتمل النجاح = $ح_١ + ح_٢$ فـــاذا افترضنا تقدم ١٠ أشخـــاص من بيروت و٨ من الشمال و٢ من الجنوب و٧ من الجبل و٩ من البقاع للحصول على وظيفة ما وقررت لجنة الاختبار سحب اسم واحد من أسماء المتقدمين يكون :

احتمال ظهور شخص من بيروت = $\frac{١٠}{١٠ + ٨ + ٢ + ٧ + ٩} = \frac{١٠}{٣٦}$

واحتمال ظهور شخص من الشمال = $\frac{٨}{٣٦}$

واحتمال ظهور شخص من الجنوب = $\frac{٢}{٣٦}$

واحتمال ظهور شخص من الجبل = $\frac{٧}{٣٦}$

واحتمال ظهور شخص من البقاع = $\frac{٩}{٣٦}$

أما احتمال ظهور شخص من خارج بيروت = $\frac{٨}{٣٦} + \frac{٢}{٣٦} + \frac{٧}{٣٦} + \frac{٩}{٣٦}$

واحتمال ظهور شخص من الجنوب والشمال = $\frac{٨}{٣٦} + \frac{٢}{٣٦}$ وهكذا

(يلاحظ ان جمع الاحتمالات يقوم على أساس ان وقوع أي حدث من الاحداث التي يحتمل وقوعها يكون مانعاً لوقوع الاحداث الاخرى ، أمـا إذا كان من الممكن أن يقع الحدثين في نفس الوقت فـلا يجوز مطلقاً جمع احتماليهما) .

ولزيادة الايضاح نأخذ مثالاً آخر ، إذا قذفنا زهرة نرد فـــان احتمال ظهور أي وجه لها يكون مانعاً لظهور الاوجه الاخرى ، وبذلك يكون احتمال ظهور الواحد أو الخمسة يساوي احتمال وقوع الواحد + احتمال وقوع الخمسة $= \frac{١}{٦} + \frac{١}{٦} = \frac{١}{٣}$ واحتمال وقوع الواحد أو الخمسة أو الثلاثة $= \frac{١}{٦} + \frac{١}{٦} + \frac{١}{٦} = \frac{١}{٢}$.

وفي هذا المثال يكون حسابنا صحيح حيث يتوفر الشرط الاساسي لجمع الاحتمالات وهو ان ظهور كل حدث يكون مـــانعاً لظهور الاحداث الاخــرى .

أما إذا رمينا قطعة من النقود إلى أعلى مرتين متتاليتين يكون احتمال ظهور الكتابة في المرة الأولى = ٢/١ وكذلك في المرة الثانية يكون احتمال ظهور الكتابة ٢/١ ، فلو كان من الجائز جمع الاحتمالين لكان احتمال ظهور الكتابة في احدى المرتين = ٢/١ + ٢/١ = ١ ومعنى ذلك اننا نكون متأكدين من ظهور الكتابة في احدى المرتين وهذا خطأ طبعاً لأننا نعلم جيدا انه قد نرمي قطعة النقود عدة مرات ولا نحصل على الكتابة في أي مرة منها والسبب في وقوعنا في هذا الخطأ هو تجاهل شرط التمانع بين الحدثين الذي يعتبر شرطاً أساسياً في قاعدة الجمع .

ضرب الاحتمالات :

ان احتمال النجاح المزدوج لحدثين مستقلين عن بعضهما يساوي حاصل ضرب احتمال الحدث الاول في احتمال الحدث الثاني ، وبديهي اننا في هذه الحالة لا نشترط أن يكون احد الحدثين مانعاً للآخر كما فعلنا في حالة جمع

الاحتمالات حيث اننا نبحث هنا في احتمال وقوع الحدثين معاً . كذلك إذا كان هناك ثلاث احداث مستقلة عن بعضها واحتمالاتها $ح_١$ ، $ح_٢$ ، $ح_٣$ على الترتيب فان احتمال حصولها جميعاً في وقت واحد = $ح_١ \times ح_٢ \times ح_٣$. وهكذا مهما تعددت الاحداث يكون احتمال حصولها جميعاً يساوي حصل ضرب احتمالاتها في بعضها . ويمكننا أن نلاحظ انه لما كان احتمال كل حدث أقل من الواحد الصحيح فان ضرب الاحتمالات يؤدي إلى جعل الاحتمال الناتج أصغر ، وهذا بالطبع أمر معقول حيث ان احتمال حصول حدثين معاً أقل من احتمال وقوع حدث واحد فقط .

مثـــال :

إذا رمينا زهرتـــين نرد فما هو احتمال أن نحصل على ستة من كل منهما سوياً .

احتمال النجاح في الزهرة الاولى $= \frac{١}{٦}$

احتمال النجاح في الزهرة الثانية $= \frac{١}{٦}$

احتمال النحــــاح المزدوج $= \frac{١}{٦} \times \frac{١}{٦} = \frac{١}{٣٦}$

المعنى الاحصائي للاحتمال :

من المناقشة السابقة يتبين لنا كيف يمكن قياس الاحتمال رياضياً فنقول مثلاً ان احتمال ظهور الستة عند رمي زهرة واحدة $= \frac{١}{٦}$ ، غير ـه بالتأمل

في هذا المقياس نجد انه يوقعنا في حيرة شديدة ذلك ان الحدث اما ان يقع أو لا يقع ولا يمكن أن يقع بخمسه أو سدسه . فعند رمي زهرة النرد فاما أن يظهر الوجه ستة او لا يظهر وبذلك بعد انتهاء التجربة نجد أن الواقع يخالف الاحتمال سواء كناقد قدرنا هذا الاحتمال ب $\frac{١}{٢}$ أو $\frac{١}{٦}$ أوأي رقم آخر قبل اجراء التجربة فالنتيجة واحدة بعد اجراء التجربة وهي اما ان يقع الحدث أو لا يقع بغض النظر عن التقدير الذي كناقد أعطينا للاحتمال من قبل.

أما المعنى الاحصائي للاحتمال فانـــه يعطينا فكرة اكثر وضوحاً من المقياس الرياضي . فالاحتمال بالمعنى الاحصائي هو تكرار نسبي فاذا كان الاحتمال الرياضي = $\frac{١}{٢}$ فمعنى ذلك احصائياً اننا اذا كررنا التجربة عدد كبير من المرات (بقرب من ∞) فان الحدث يقع في نصف هـذه المرات ولا يقع في النصف الآخر . فاذا رمينا قطعة نقود ١٠٠٠٠٠ مرة فان النتيجة لا يمكن أن تختلف كثيراً عن ٥٠٠٠٠ وجه و ٥٠٠٠٠ ظهر الا اذا كانت القطعة ليست كاملة التوازن .

على ضوء هذا المعنى نعود للتوزيع التكراري لعدد مرات النجاح الذي سبق الاشارة اليه ولايجاد متوسط عدد مرات النجاح نضرب كل قيمة (من قيم عدد مرات النجاح) في الاحتمال الخاص بها ونقسم على مجموع التكرارات (مجموع الاحتمالات) وهو يساوي ١ فينتج ان متوسط قيم مرات النجـاح يساوي ن × ح أي عدد المحاولات مضروباً في احتمال الكسب في أية محاولة باعتبار ان نجاح أي محاولة لا يؤثر في احتمال نجـــاح أو فشل أي محاولة أخرى وبفرض ان احتمال النجاح واحد في كل المحاولات .

مثـال :

رمى لاعب ١٢ زهرة دفعة واحدة وكرر هذه التجربة عدة مرات فما هو متوسط عدد ما يحصل عليه من الشيشات لكل مرة .

احتمال رمي الشيش من أي زهرة = $\frac{١}{٦}$

ورمي ١٢ زهرة دفعة واحدة يعني عمل ١٢ محاولة .

متوسط عدد ما يحصل عليه من الشيشات في كل رمية = ١٢ × $\frac{١}{٦}$ = ٢

والمقصود بقولنا ان الوسط الحسابي لعدد الزهرات الناجحة = ٢ هو اننا اذا قمنا فعلاً بتجربة قد نحصل على اكثر أو أقل من زهرتين وربما نحصل على اثنتين بالضبط ولو كررنا التجربة عدداً كبيراً من المرات فان المتوسط يكون ٢ أو قريباً منها .

وكما أوجدنا الوسط الحسابي للتوزيع التكراري لعدد مرات النجاح يمكننا أن نوجد الانحراف المعياري لهذا التوزيع وهو يقيس لنا درجة الاختلاف بين عدد مرات النجاح التي يمكن ان نحصل عليها والوسط الحسابي . ويمكن ايجاد الانحراف بضرب مربع كل قيمة من قيم عدد مرات النجاح في الاحتمال الخاص بها ثم نطرح مربع الوسط الحسابي من مجموع حواصل الضرب فينتج مربع الانحراف المعياري (التباين) ويمكن اثبات ان الانحراف المعياري للتوزيع التكراري لعدد مرات النجاح = $\sqrt{\text{ن ح ل}}$

مثـال :

إذا اعتبر ظهور جهار أو أكثر من زهرة النرد نجاحاً فما هو متوسط عدد الزهرات الناجحة إذا رمينا ١٢ زهرة ومـــا هو الانحراف المعياري من هذا المتوسط .

احتمال النجاح في هذه الحالة بالنسبة لكل زهرة = $\frac{٣}{٦}$ = $\frac{١}{٢}$

متوسط عدد ما نحصل عليه من زهرات ناجحة في كل مرة =

$$١٢ \times \frac{١}{٢} = ٦ \text{ زهرة}$$

وحيث أن احتمال النجاح = $\frac{١}{٢}$ واحتمال الفشل = ١ − $\frac{١}{٢}$ = $\frac{١}{٢}$

$$\therefore ع = \sqrt{١٢ \times \frac{١}{٢} \times \frac{١}{٢}} = \sqrt{٣} = ١٫٧٣٢ \text{ تقريباً}$$

مثـال :

رميت قطعة نقود كاملة التوازن ٩٠٠ مرة فالنتيجة التي نتوقعها نظرياً هي أن نحصل على وجه القطعة نصف عدد المرات تقريباً ونصل إلى هـذه النتيجة بايجاد الوسط الحسابي لعدد مرات النجاح (يعتبر ظهور وجه قطعة النقود نجاحاً) وهو في هذا المثال = ٩٠٠ × ٢/١ (احتمال النجاح) = ٤٥٠ نجاحاً في كل مرة نعيد فيها التجربة فعوامل المصادفة يمكن أن تجعل النتائج تختلف قليلاً من ذلك فقد نحصل على ٤٥٥ نجاحاً أو ٤٤٨ أو ٤٤٦ أو ٤٥٢ وهكذا . الا اننا نستطيع أن نقول اننا إذا أعدنا التجربة عدداً كبيراً جداً من المرات فان المتوسط لعدد مرات النجاح في كل التجارب التي نجريها سوف يكون ٤٥٠ .

الا أننا لا نستطيع أن نؤكد أن الفرق بين عدد مرات النجاح الواقعي والعدد المتوقع (الوسط الحسابي) يكون دائماً راجعاً الى عوامل الصدفة ، فقد تكون قطعة النقود ليست كاملة التوازن وبذلك يكون هناك تحيزاً اما نحو وجه القطعة أو ظهرهـــا ، كذلك يمكن أن يتسرب عنصر التحيز الى طريقة رمي قطعة النقود. وبالطبع لو كان ليس بامكاننا أن نميز بين الأخطاء (الفرق بين الرقم الواقعي والرقم المتوقع) التي يمكن أن تكون راجعة الى الصدفة وتلك التي ترجع الىعوامل اخرى كالتحيز أو أي عامل آخر يؤثر في

نتيجة التجربة فان مانحصل عليه من نتائج يكون معدوم القيمة حيث يكون دائماً موضع الشك والتساؤل، هل تختلف النتيجة التي نحصل عليها عن النتيجة المتوقعة اختلافاً يمكن أن نعزوه الى الصدفة أو هو اختلاف جوهري راجع الى عوامل اخرى ، يمكن أن يكون التحيز احداها .

على انه احصائياً يمكن التمييز بين الفروق التي تكون راجعة الى عوامل الصدفة وتلك التي تكون راجعة الى عوامل اخرى وذلك باستخدام خواص التوزيع التكراري لعدد مرات النجاح . إذ لو كررنا المحاولة عدداً كبيراً جداً فان نتائج هذه المحاولات تتوزع توزيعاً معتدلاً تقريباً متوسطة = ن ح وانحرافه المعياري $\sqrt{\text{ن ح ل}}$. وبذلك يلعب الانحراف المعياري للتوزيع دوراً هاماً في تحديد ما اذا كان الفرق راجعاً الى الصدفة أو غيرها . فاذا كان الفرق بين النتيجة التي نحصل عليها فعلاً والنتيجة المتوقعة أقل من الانحراف المعياري $\sqrt{\text{ن ح ل}}$ فمن المحتمل أن يكون هذا الفرق راجعاً إلى عوامل الصدفة ، أما اذا كان الفرق يعادل أو يزيد عن ضعف الانحراف المعياري فان احتمال حدوث الفرق بطريقة المصادفة يكون ضئيلا ، وكذلك اذا كان الفرق يعادل أو يزيد عن ثلاث أمثال الانحراف المعياري يكاد يكون من المستحيل عملياً أن يكون راجعاً الى عوامل الصدفة. على أن هناك خلاف بين الاحصائيين فيما يتعلق بالتمييز بين الفروق الظاهرية (الراجعة الى الصدفة) والفروق الجوهرية (التي ترجع الى عوامل غير الصدفة) ، بعضهم يعتبر الفرق الذي يعادل ٢ $\sqrt{\text{ن ح ل}}$ فرقاً جوهرياً لا يمكن أن يعزى الى عوامل الصدفة والبعض الآخر يشترط لاعتبار الفرق جوهرياً أن يكون مساوياً ٣ $\sqrt{\text{ن ح ل}}$ على الأقل .

مثال ٤٣ :

انتجت احدى مزارع البرتقال ١٠٠٠٠ ثمرة وجد بينها ٥٧٠ ثمرة عطبة

بينما كانت النسبة المعتادة للعطب هي ١٠٪ هل يدل الفرق بين عدد الثمار العطبة الذي نتوقعه نظرياً وذلك الذي حصلنا عليه فعلاً على حدوث تغير جوهري في ظروف الزراعة أم يمكن أن نعزوه الى عوامل المصادفة .

$$\text{نسبة الثمار العطبة} = \frac{٧٥٠}{١٠٠٠٠} \times ١٠٠ = ٧٫٥ ٪$$

والمطلوب معرفة هل الفرق بين ٧ر٥ ٪ وهي النسبة الواقعية و ١٠ ٪ وهي للنسبة المتوقعة فرق ظاهري أو فرق جوهري .

$$\text{عدد الثمار التي يتوقع عطبها} = ١٠٠٠٠ \times \frac{١}{١٠} = ١٠٠٠ \text{ ثمرة}$$ (الوسط الحسابي) .

$$\text{الانحراف المعياري} = \sqrt{١٠٠٠٠ \times \frac{١}{١٠} \times \frac{٩}{١٠}} = ٣٠$$

الفرق بين عدد الثمار العطبة الواقعي والمتوقع = ١٠٠٠–٥٧٠=٤٣٠ ثمرة.

$$\text{هذا الفرق} = \frac{٤٣٠}{٣٠} \text{ ع} = ١٤٫٣٣ \text{ ع}$$

وهو فرق يزيد كثيراً على ثلاث أمثال الانحراف المعياري فهو اذن فرق جوهري يدل على تحسن فعلي في ظروف الزراعة .

الوسط الحسابي والانحراف المعياري لنسبة النجاح :

في الفقرات السابقة ناقشنا الوسط الحسابي والانحراف المعياري لعدد مرات النجاح ، غير أنه من الأفضل احياناً أن نعبر عن مرات النجاح في شكل

نسبة مئوية بدلاً من التعبير عنها في شكل اعداد مطلقة . وفي هـذه الحالة

يكون الوسط الحسابي لنسبة النجاح $= \frac{ن ح}{ن} =$ ح والانحـراف المعياري

للنسبة $= \sqrt{\frac{ن ح ل}{ن^٢}}$ أي $\sqrt{\frac{ح \times ل}{ن}}$ وعند استعمال هذه القوانين يجب أن

نتنبه إلى أن ح و ل فيها ليست نسب مئوية بـل نسب منسوبة الى الواحد الصحيح . وعلى أساس هذه القوانين يمكن حل المثال السابق كالآتي :

متوسط نسبة النجاح $= \frac{١}{١٠}$ أي ١٠ ٪

$$ع = \sqrt{\frac{\frac{١}{١٠} \times \frac{٩}{١٠}}{١٠٠٠٠}} = ٠٠٣ر$$

ع للنسبة المئوية = ٠٠٣ر × ١٠٠ = ٣ر

الفرق بين النسبة المتوقعة والنسبة الفعلية بدلالة ع $= \frac{٥ر٧٠١٠}{٣ر} =$ ١٤ر٣٣

وهو فرق أكبر بكثير من ٣ أمثال ٣ر فهو اذن فرق جوهري .

ويمكن أن نصل إلى نفس النتيجة إذا أجرينا حسابنا على النسبة الواقعية ٥ر٧ ٪ .

$$\text{حيث} \quad ع = \sqrt{\frac{\frac{٥٧٠}{١٠٠٠} \times \frac{٤٣٠}{١٠٠٠}}{١٠٠٠٠}} = ٠٠٤٩ر \ .$$

ع للنسبة المئوية = ٠٠٤٩ر × ١٠٠ = ٤٩ر

الفرق بين النسبة الواقعية والمتوقعة بدلالة ع = $\frac{٥٫٨٠١٠}{٠٫٠٤٩}$ = ٧٫٧

وهو فرق أكبر بكثير من ٣ أمثال ٠٫٠٤٩ فهو اذن فرق جوهري .

مثـــال ٤٤ :

انبت احد أنواع البسلة ١٤٣٠٠ من البذور الصفراء و٤٩٠٠ من البذور الخضراء في حين ان عدد البذور الصفراء الذي نتوقعه حسب قانون مندل يعادل ٧٥ ٪ من مجموع البذور ، هل يمكن أن نعزو الفرق بين العدد الذي حصلنا عليه فعلا وذلك الذي نتوقعه نظرياً إلى عوامل المصادفة :

عدد البذور الصفراء المتوقعة = ١٩٢٠٠ × $\frac{٣}{٤}$ = ١٤٤٠٠

الفرق بين العدد المتوقع والعدد الفعلي = ١٤٤٠٠ − ١٤٣٠٠ = ١٠٠

ع = $\sqrt{ن ح ل}$ = $\sqrt{١٩٢٠٠ \times \frac{٣}{٤} \times \frac{١}{٤}}$ = $\sqrt{٣٦٠٠}$ = ٦٠

الفرق بدلالة ع = $\frac{١٠٠}{٦٠}$ = ١٫٦٦

وهو أقل من ٣ أمثال ع فهو اذن فرق ظاهري يمكن أن نعزوه إلى المصادفة أي ان النتيجة لا تعتبر مخالفة لقانون مندل .

ويمكن اجراء العمل بالطريقة الآتية :

ع = $\sqrt{\frac{\frac{٣}{٤} \times \frac{١}{٤}}{١٩٢٠٠}}$ = ٠٫٠٠٣١

ع للنسبة المئـوية = ٠٫٠٠٣١ × ١٠٠ = ٠٫٣١

$$\text{الفرق بين النسبتين} = ٧٥ - \frac{١٤٣٠٠}{١٩٢٠٠} \times ١٠٠ = ٥$$

هذا الفرق أقل بكثير من ٣ امثال ع (٩٢ر) ، وهو بذلك فرق يمكن أن يكون راجعاً إلى الصدفة

ويمكن اجراء العمل بالطريقة الآتية :

$$ع = \sqrt{\frac{\frac{١٤٣}{١٩٢} \times \frac{٤٩}{١٩٢}}{١٩٢٠٠}} = ٠٠٣١ر$$

ع للنسبة المئوية = ٠٠٣١ر × ١٠٠ = ٣١ر

$$\text{الفرق بين النسبتين } ٧٥ - ٧٤ر٥ = ٥ر \text{ بدلالة } ع = \frac{٥ر}{٣ر} = ١ر٦٦$$

أي أقل من ٣ ع فهو اذن فرق يمكن أن يكون راجعاً إلى الصدفة .

تطبيق هذه النتائج في دراسة العينات :

عندما نسحب عينة من مجتمع ما بواسطة الارقام العشوائية فانها تكون احدى عينات كثيرة (او محـــاولات كثيرة) يمكن أن تسحب من هذا المجتمع وإذا كان موضوع دراسة العينة هو احتمال ظاهرة ما (احتمال البطالة أو احتمال الوفاة او احتمال النجاح الخ ..) فاننا نستطيع أن نكوّن توزيعاً تكرارياً لعدد العينات التي تعطي كل احتمال من احتمالات الظاهرة موضوع الدرس وقد رأينا ان التوزيـع التكراري لهذه الاحتمالات المختلفة عندمـــا يكون عدد المحاولات كبيراً جداً (عدد العينات) يتخذ شكل

منحنى طبيعي ويكون الانحراف المعياري لهذا التوزيع بدلالة الاحتمال مساوياً $\sqrt{ن ح ل}$ وبشكل آخر $\sqrt{\frac{ح ل}{ن}}$ إذا أردنا قياس الانحراف المعياري لاحتمال النسبة . وعلى هذا الأساس نستطيع أن نحسب مدى الانحراف بين الاحتمال المقدر من العينة والاحتمال العام للمجتمع الاصلي الذي يمثله متوسط توزيع الاحتمالات الخاصة بالعينات المختلفة المسحوبة من هذا المجتمع .

وفي اجراء الاختبارات المختلفة بالنسبة للظواهر الوصفية على أساس العينات سوف نتبع نفس الخطوات التي سبق أن أشرنا اليها بالنسبة للظواهر الرقمية (المتوسطات) وبذلك لا يكون هناك داعي لشرح خطوات العمل ونكتفي لذلك باعطاء امثلة مع نماذج حلولها .

تقدير نسبة ظاهرة ما في مجتمع معين من واقع بيانات العينة :

مثال ٤٥ : اجريت دراسة بالعينة تبين منها ان ٧٠٠شخص مرضى بمرض معين من مجموع عدد الأشخاص في العينة والبالغ عددهم ٢٥٠٠ شخص . المطلوب تقدير نسبة الاصابة بهذا المرض في المجتمع .

$$\text{نسبة الاصابة في العينة} = \frac{٧٠٠}{٢٥٠٠} \times ١٠٠ = ٢٨ \%$$

$$\text{الخطأ المعياري لنسبة الاصابة (للوحدة الواحدة)} = \sqrt{\frac{ح ل}{ن}} = \sqrt{\frac{٠٫٢٨ \times ٠٫٧٢}{٢٥٠٠}}$$

$$= ٠٫٠٠٨٩$$

الخطأ المعياري للنسبة المئوية للاصابة = ٠٫٠٠٨٩ × ١٠٠ = ٫٨٩

∴ فترة الثقة بدرجة فئة ٩٥ ٪ = ± ٫٨٩ × ١٫٩٦ = ١٫٧٤

∴ نسبة الاصابة في المجتمع بدرجة ثقـــة ٩٥ ٪ تتراوح بين ٢٨ ٪ ± ١٫٧٤ ٪ = ٢٩٫٧٤ ، ٢٦٫٢٦

اختبار معنوية الفرق بين نسبة ظاهرة ما في مجتمع معين وفي العينة :

مثال ٤٦ : في عام ١٩٤٢ كان عدد المواليد في الجمهورية العربية المتحدة ٦٥٨٣٢٤ منهم ٣٤٣٦٣٣ ذكر وفي نفس العام كان عدد المواليد في احدى المحافظات ٣٢٢٧٧ منهم ١٦٦٨٣ ذكر – فهـل يستدل من ذلك على وجود اختلاف جوهري في نسبة الذكور من المواليد ؟

احتمال أن يكون المولد ذكراً في القطر كله = $\frac{٣٤٣٦٣٣}{٦٥٨٣٢٤}$

= ٠٫٥٢٢ تقريباً

احتمال أن يكون المولد ذكراً في المحافظة = $\frac{١٦٦٨٣}{٣٢٢٧٧}$

= ٠٫٥١٦ تقريباً

الخطأ المعيار لاحتمال ان يكون المـــولود ذكراً =

$\sqrt{\frac{٠٫٥١٦ \times ٠٫٤٨٤}{٣٢٢٧٧}}$ = ٫٠٠٢٧

الفرق بدلالة ع = $\frac{٠٫٥٢٢ - ٠٫٥١٦}{٫٠٠٢٧}$ = ٢٫٢ تقريباً

وبذلك يكون الفرق واقعاً ضمن درجة ثقة ٩٩ ٪

∴ يمكن أن يكون هذا الفرق راجعاً الى الصدفة .

ويمكن الوصول إلى نفس النتيجة بطريقة اخرى حيث يكون عدد المواليد الذكور المتوقع في الجيزة على اساس الاحتمال العام = ٣٢٢٧٧ × ٠ر٥٢٢ = ١٧٨٤٩ تقريباً .

وبذلك يكون الفرق بين العدد المتوقع والعد الحقيقي =

١٧٨٤٩ − ١٦٦٨٣ = ١٦٦

الخطأ المعياري لعـــدد المواليد المتوقع =

$\sqrt{٣٢٢٧٧ \times ٠ر٥٢٢ \times ٠ر٤٧٨}$ = ٨٩ تقريباً

الفرق بدلالة ع = $\frac{١٦٦}{٨٩}$ = ١ر٨

وبذلك يكون الفرق راجعاً الى الصدفة .

اختبار معنوية الفرق بين عينتين :

مثال ٤٧ :

في عينة من ٣٠٠ أسرة كانت نسبة الاسر التي يزيد دخلهـــا عن ١٥ دينار هي ٣٥ ٪ وفي عينة أخرى من ٢٥٠ أسرة كانت النسبة ٣٠ ٪ — اختبر معنوية الفرق بين العينين : —

الفرق الواقعي بين النسبتين = ٣٥ − ٣٠ = ٥ ٪

$$\text{ع ف} = \sqrt{\frac{٠ر٣٥ \times ٠ر٦٥}{٣٠٠} + \frac{٠ر٣ \times ٠ر٧}{٢٥٠}}$$

$$= ٠ر٠٤$$

ع ف للنسبة المئوية = ٠ر٠٤ × ١٠٠ = ٤

بدرجة ثقة ٩٥ ٪ يكون مدى الثقة صفر $\pm$ ٥ × ١ر٩٦ + ٧ر٨٤

وحيث أن الفرق ٥٪ أقل من هذا الفرق المحتمل ، اذا الفرق بين العينتين يمكن أن يكون راجعاً الى الصدفة .

حل آخر

نفترض أن العينين تنتميان لمجتمعين متشابهين من حيث نسبة الأسر التي يزيد دخلها عن ١٥ دينار ، أي لها نفس نسبة هذه الأسر ، وحيث أن هذه النسبة العامة لهذين المجتمعين غير معروفة فاننا نقدرها كوسط حسابي مرجح للنسبتين في العينين .

$$\text{أي} \quad \frac{٠ر٣٥ \times ٣٠٠ + ٣٠ \times ٢٥٠}{٥٥٠} = ٠ر٣٣$$

ويكون الخطأ المعياري لمجتمع الفروق بين العينات العشوائية التي يمكن سحبها من كلا المجتمعين $\text{ع ف} = \sqrt{٠ر٣٣ \times ٠ر٦٧ \left(\frac{١}{٣٠٠} + \frac{١}{٢٥٠}\right)}$

$$= ٠ر٠٤$$

وبذلك يكون الخطأ المعياري للنسبة المئوية بدرجة ثقة ٩٥٪ .

$$= \pm ٠ر٠٤ \times ١٠٠ \times ١ر١٩٦ = ٨ر٧٤$$

وحيث ان الفرق بين العينتين اقل من هذا الفرق المحتمل ، اذا الفرق بين العينتين يمكن ان يكون راجعاً الى الصدفة .

لوحات مراقبة الانتاج :

اشرنا سابقاً الى كيفية اجراء المراقبة الاحصائية للانتاج على اساس تصميم لوحات تبين عليها فترات الثقة المطلوبة ثم يجري الاختبار على وحدة واحدة من الانتاج ويعين موقعها في لوحة المراقبة ، ثم بعد فترة من الزمن نؤخذ وحدة اخرى ونبين موقعها في لوحة المراقبة وهكذا يجري العمل على اساس وحدات كل وحدة منها قائمة بذاتها . وحيث انه في الغالب تجري المراقبة ليس على وحدات ، كل وحدة على حدة ، وانما على اساس عينات من الانتاج فيجدر بنا ان ندرس كيف يتم العمل في هذه الحالة .

وفكرة المراقبة في هذه الحالة لا تختلف بتاتاً عن الفكرة التي سبق الاشارة اليها ، حيث ان جميع العينات التي يحتمل أخذها من الانتاج تكون توزيعاً معتدلاً يكون متوسطه هو المتوسط العام للظاهرة موضوع المراقبة ونسبته هي النسبة العامة للظاهرة موضوع المراقبة . واذا كانت الفروق بين العينات والمجتمع الذي سحبت منه ناتجة عن عوامل الصدفة تكون مقاييسها واقعة ضمن فترات الثقة التي حددناها على لوحة المراقبة ، أما إذا كانت خارجة عن هذه الفترات فيكون الفرق بينها وبين المجتمع الذي سحبت منه فرقاً جوهرياً ناتجاً عن أي عامل آخر غير الصدفة ، وقد يكون خللاً في آلات الانتاج أو اهمالاً من العامل أو عدم انتظام الآلات في عملها . ومن الواضح أن الانحراف المعياري لتوزيع العينات (الخطأ المعياري) يحسب بالقانون

$\sqrt{\frac{ع}{ن}}$ إذا كنا بصدد درس متوسط ظاهرة معينة أو بالقانون $\sqrt{\frac{ح \times ل}{ن}}$

اذا كنا بصدد درس نسبة ظاهرة معينة .

مثال ٤٨ :

يستخدم مصنع ما آلة تنتج انابيب متوسط قطرهـا ٥٧٤ر٠ بوصة بانحراف معياري ٠٠٨ر٠ بوصة . صمم لوحة لمراقبة انتاج هذه الآلة على أساس عينات تتكون كل منها من ٦ انابيب كل ساعتين عمل .

$$\text{الخطأ المعياري} = \sqrt{\frac{٠٥ر٠ \times ٩٥ر٠}{٢٠٠}} = ٠١٥٤ر٠$$

بدرجة ثقة ٩٩ ٪ يكون حدى المراقبة بين

$٥٪ \pm ١٠٠٥٤ر٠ \times ١٠٠ \times ٥٨ر٢$

$= ٥٪ \pm ٩٧٣٢ر٣$

أي بين ٩٧٣٢ر٨ ، ١٠٦٨ر١

الحد الأعلى للمراقبة	....	٩٪ تقريباً
المواصفة	....	٥٪
الحد الأدنى للمراقبة	....	١٪ تقريباً

واضح من الرسم ان النقطة الاخيرة تدل على ان نسبة التلف قد زادت الى

درجة لا يمكن ان يكون الفرق عندها راجعة الى الصدفة ، فلا بد لذلك من الكشف على اسباب هذا الفرق الجوهري .

تحديد حجم العينة :

ان تحديد حجم العينة في بحث ما امر هام جداً حيث يتوقف عليه تحديد التكاليف التي يتطلبها البحث والمجهودات التي سوف تبذل فيه . ولا يمكن اجراء هذا التحديد الا بالاجابة على الاسئلة الآتية :

١ – ما هو الخطأ الذي نكون مستعدين لقبوله في البحث ؟ ولا شك ان ذلك يتوقف بالدرجة الاولى على الغرض الذي يجري من اجله البحث أي على الناحية التي سوف تستخدم فيها النتائج . ومن الواضح ان بعض البحوث تتطلب دقة كبيرة الا ان ذلك يعني ان حجم العينة لا بد ان يكون كبيراً ، الأمر الذي يؤدي الى رفع التكاليف ، وبعض البحوث الاخرى لا تتطلب ذلك وبذلكِ يمكن قبول خطأ أكبر لانقاص التكاليف والجهود .

٢ – ما هي درجة الثقة التي نرغب في تحقيقها من ناحية الخطأ ؟ ومن الواضح ان زيادة درجة الثقة تتطلب حجما كبيراً للعينة ؛ والعكس صحيح ، وبذلك يرتبط الامر ايضاً بتكاليف البحث والجهود التي يمكن توفيرها له .

٣ – لا بد من معرفة معالم المجتمع الذي نريد أن نأخذ عينة منه ، وتتوقف هذه المعالم على نوع الدراسة المطلوب إجرائها ، فاذا كنا بصدد بحث عن متوسط ظاهرة ما فاننا نحتاج بذلك أن نتعرف على تباين المجتمع بالنسبة لهذه الظاهرة ، وإذا كنا بصدد بحث عن نسبة ظاهرة ما لا بد أن يكون لدينا معرفة تقريبية من هذه النسبة .

٤ – إذا كنا ندرس عينة طبقية فانه يمكن حساب حجم العينة لكل قسم من المجتمع وبالجمع يمكن حساب حجم العينة الكلي .

٥ – يمكن أن يكون البحث بالعينة لدراسة أكثر من ظاهرة واحدة في المجتمع ، وحجم العينة الذي يصلح لدراسة ظاهرة منها قد لا يكون صالحاً لدراسة باقي الظواهر . وفي هذه الحالة نحدد الأخطاء حسب الظواهر الأكثر أهمية في الدراسة ونحدد حجم العينة الذي يناسب كل من هذه الظواهر المنفصلة فاذا كانت الاحجام متقاربة فنأخذ أكبر حجم منها طالما كان في حدود النفقات والجهود التي نستطيع أن نتحملها . أما اذا كانت الاحجام متفاوتة كثيراً فلا بد أن نضحي ببعض الدقة بالنسبة لبعض الظواهر حتى نجعلها متقاربة وتأخذ الحجم الأكبر . أما إذا كان التفاوت من الكبر بحيث لا يمكن مع تضحيه بعض الدقة ان نجعلها متقاربة ففي هذه الحالة لا بد أن نهمل بعض الظواهر من البحث .

فاذا استطعنا أن نعرف معالم المجتمع وأن نحدد درجة الخطأ التي نقبلها ودرجة الثقة في هذا الخطأ فان تحديد حجم العينة يكون مجرد استنتاج جبري من إحدى معادلات الخطأ المعياري (تبعاً لنوع الدراسة) . والخطأ المقبول قد يكون في صيغة خطأ مطلق أو في صيغة معامل اختلاف لهذا الخطأ أي في صيغة نسبة مئوية من التقدير المطلوب .

مثال ٥٠ :

في مجتمع حجمه ٤٣٠ وحدة وجد ان متوسطه = ١٩ وتباينه = ٨٥ر٦
أوجد عدد الوحدات التي يجب سحبها في عينة عشوائية لتقدير المتوسط بحيث يكون الخطأ ١٠ ٪ من التقدير باحتمال ٩٥ ٪ .

حيث ان الخطأ معطى في شكل معامل اختلاف وحيث ان درجة الثقة المطلوبة هي ٩٥ ٪ يكون قانون الخطأ المعياري الذي نستعمله هو :

$$\text{ع} = \frac{\dfrac{\text{ع}}{\sqrt{\text{د}}}\,٢}{\bar{\text{س}}} \times ١٠٠$$ (استخدمنا ٢ تقريباً)

$$\frac{٢\sqrt{٨٥٫٦}}{\sqrt{\text{د}}}$$

$$\therefore ١٠ = \frac{٢\sqrt{\dfrac{٨٥}{\text{د}}}}{١٩} \times ١٠٠$$

بالتربيع $١٠٠ = ٤ \times \frac{٨٥٦}{١٠} \times \frac{١}{\text{د}} \times \frac{١}{١٩} \times \frac{١}{١٩} \times ١٠٠ \times ١٠٠$

$$١ = \frac{٣٤٢٤٠٠}{٣٦١٠\,\text{د}}$$

∴ ن = ٩٥ تقريباً .

(نلاحظ انه حيث ان حجم المجتمع محدد ، كان من الأفضل من ناحية لدقة استخدام القانون $\text{ع} = \frac{\text{ع}}{\sqrt{\text{ن}}}\sqrt{\frac{\text{ن} - \text{د}}{\text{ن} - ١}}$

(حيث ن حجم المجتمع، د حجم العينة) إلا أن الفرق في النتيجة لا يكون من الأهمية بمكان كبير) .

مثـــال ٥١ :

إذا كان متوسط الدخل في المجتمع ٢٥ ديناراً ، وتباينة ٢٢٥ ديناراً .

كم يكون حجم العينة اللازمة للحصول على معـــامل اختلاف للمتوسط ١٠ ٪ .

$$ع = \frac{\sqrt{\frac{ع}{ن}}}{\bar{س}} \times ١٠٠$$

$$١٠٠ = \frac{\sqrt{\frac{ع}{ن}}}{٢٥} \times ١٠٠$$

بالتربيع $١٠٠ = \frac{٢٢٥}{ن} \times \frac{١}{٦٢٥} \times ١٠٠٠٠$

$\therefore$ ٦٢٥ ن = ٢٢٥٠٠

$\therefore$ ن = ٣٦ وحدة .

مثال ٥٢ :

كم يكون حجم العينة لدراسة نسبة انتشار مرض ما اذا كان من المعروف ان هذا المرض ينتشر في العادة بنسبة ١ ٪ تقريباً وإذا كان الخطأ المطلوب هو ٢٠ ٪ من التقدير . كم يكون حجم العينة اذا كان الخطأ المطلوب هو بالنسبة لكل من الذكور والاناث .

حيث ان الدراسة خاصة بنسبة ظاهرة ما نستخدم قانون الخطأ المعياري للنسبة في شكل معامل اختلاف حيث ان الخطأ معطى في شكل نسبة مئوية من التقدير .

$$ع = \frac{\sqrt{\frac{ح(١ - ح)}{ن}}}{ح} \times ١٠٠ \quad (\text{حيث } ل = ١ - ح)$$

$$\therefore ٢٠ = \frac{\sqrt{\frac{٠١ر \times ٩٩ر}{ن}}}{١ر} \times ١٠٠$$

بالتربيع

$$٤٠٠ = \frac{١}{١٠٠} \times \frac{٩٩}{١٠٠} \times \frac{١}{ن} \times \frac{١٠٠}{١} \times \frac{١٠٠}{١} \times ١٠٠٠٠$$

$$\therefore ٤ = \frac{٩٩٠٠}{ن}$$

$$\therefore ٤ن = ٩٩٠٠ \qquad \therefore ن \; ٢٤٧٥$$

ونفس الحجم نحتاجه لكل من الذكور والاناث أي أن العينة في مجموعها يجب أن تتكون من ٤٩٥٠ نصفهم من الذكور والنصف الآخر من الاناث .

وواضح اننا اذا أردنا خطأ أقل بنسبة ١٠ ٪ فقط مثلاً فان حجم العينة لكل من الذكور والاناث يجب أن يكون ٩٩٠٠ وحدة أي أن العينة في مجموعها يجب أن تكون ١٩٨٠٠ وحدة . أما اذا قبلنا خطأ أكبر فان حجم العينة يقل . واذا كنا نريد نفس الخطأ السابق ٢٠٪ ولكن بدرجة ثقة ٩٥٪

فان حجم العينه يجب أن يزيد . ولهذا بعد تحديد حجم العينة يجب أن نعلم ما اذا كان هذا الحجم يتناسب مع التكاليف والجهود التي في استطاعتنا تخصيصها للدراسة أم لا ، فاذا كان يتناسب معهـــا أجرينا البحث بالحجم المحدد ، أما اذا كان يفوق هذه التكاليف والجهود علينا أن نقنع بدرجة دقة أقل أي بخطأ معياري أكبر ، أو نعمل على اجراء آخر نستطيع به أن ننقص من الانحراف المعياري للمجتمع وذلك باتباع نوع آخر من العينات حيث قـد يكون من الممكن تقسيم هـــذا المجتمع الى طبقات متجانسة وبذلك ينقص الانحراف المعياري فنستطيع بذلك انقاص حجم العينة مع الابقاء على الدقة المطلوبـــة .

ومن ناحية الدراسات الخاصـة بالنسب نلاحظ أن حجم العينة يتناسب تناسب عكسياً مع احتمال وجود الظاهرة في المجتمع ، اذا كلما كان الاحتمال صغيراً أي كلما كانت النسبة صغيرة كلما كان حجم العينة التي يجب أخذهـــا كبيراً (الشيء النادر يحتاج الى بحث أوسع وأعمق) ، والعكس كلما كانت الظاهرة شائعة في المجتمع كلما كان في استطاعتنا أن نأخذ عينة صغيرة ونحقق نفس درجة الدقة المطلوبة ، ففي المثال السابق اذا كان المرض منتشراً بنسبة ١٠٪ وليس ١٪ فقط فان حجم العينة التي يجب أخذها يكون :

$$١٠٠ \times \frac{\sqrt{\frac{١ر \times ٩ر}{ن}}}{١ر} = ٢٠$$

$$١٠٠٠٠ \times \frac{١٠}{١} \times \frac{١٠}{١} \times \frac{١}{ن} \times \frac{٩}{١٠٠} \times \frac{١}{١٠٠} = ٤٠٠ \therefore$$

$$٩٠٠ = ٤ن \therefore$$

$$\text{٢٢٥ وحدة} = ن \therefore$$

تمارين

١ – اذا علمت ان متوسط عمر اللمبة الكهربائية حسب عينة مكونة من ١٠٠ لمبة هو ١١٩٠ بانحراف معيـــاري ٩٠ ساعة ، ومن عينة اخرى مكونة من ٧٥ لمبة ظهر ان المتوسط هو ١٢٣٠ ساعة بانحراف معياري ١٢٠ ساعة . اختبر معنوية الفرق بين العينين .

٢ – ينتج احد المصانع جوارب نسبة التلف فيها ٣ ٪ – صمم لوحة لمراقبة الانتاج على أساس عينات تتكون كل منها من ٢٠٠ جورب .

٣ – ينتج احد المصانع خيطان قوة تحملها ٨٫٦٤ أوقيه بانحراف معياري ١٫٢٨ اوقية – صمم لوحة لمراقبة الانتاج على اساس عينات تتكون كل منها من ١٦ خيط .

٤ – تبين من عينة عشوائية مكونة من ١٠٠ عامل ان متوسط الأجر هو ٢٧٫٥ لـــيرة بانحراف معياري ٥ ليرات – قدر متوسط الأجر لعمال الصناعة كلها .

الفصل الحادي عشر

الارتباط

يقصد بالارتباط بين ظاهرتين أو بين متغيرين وجود علاقة بينهما بمعنى انه اذا تغير أحد المتغيرين في اتجاه معين (بالزيادة أو بالنقص) فان المتغير الآخر يميل الى التغير في اتجاه معين أيضاً . والمتغيران اما أن يتغيرا في نفس الاتجاه (بالزيادة أو بالنقص) فيسمى الارتباط طردياً ، واما أن يتغيرا في اتجاهين مختلفين فاذا اتجه أحدهما نحو الزيادة اتجه الآخر نحو النقص وفي هـذه الحالة يسمى الارتباط عكسياً .

ووجود الارتباط لا يعني ان أي تغير في أحد المتغيرين لا بد أن يكون مصحوباً بتغير في المتغير الآخر ، وانما يعني أنه في أغلب الحـالات يصحب التغير في احدى الظاهرتين تغير الظاهرة الأخرى ، وذلك هو السبب في اننا نقول أن التغير في أحد المتغيرين يصحبه ميل الظاهرة الاخرى نحو التغير .

ويقاس الارتباط بمعامل الارتباط الذي يدلنا فقط على درجة العلاقة بين الظاهرتين ، هل هي علاقة قوية أو ضعيفة ، ولكنه لا يفسر لنا السبب في هذه العلاقة . والارتباط بين ظاهرتين يمكن أن يرجع الى واحد من الاسباب الآتية : —

١ — ان احد المتغيرين هو نتيجة مباشرة للمتغير الآخر مثل الارتباط بين سعر سلعة ما والطلب عليها أو عرضها .

٢ — ان احد المتغيرين هو سبب غير مباشر للمتغير الآخر ، فمثلا يكون ارتفاع ثمن القطن في مصر سبباً غير مباشر في ارتفاع ثمن القمح ، اذ يترتب على ارتفاع ثمن القطن زيادة المساحة المزروعة قطناً فتقل المساحة المزروعة قمحاً وبذلك يقل عرضه في السوق فيرتفع ثمنه .

٣ — ان المتغيرين سوياً هما نتيجة لعامل واحد مشترك ، مثلا التغير في سعر سلعتين تستهلكهما طبقة واحدة من الناس .

٤ — وجود عامل مشترك بين العوامل المختلفة التي تؤثر على كل من المتغيرين ، مثلا الارتباط بين ذكاء الطالب في مادتين .

والمهم أن معامل الارتباط لا يحدد لنا أي من هذه الأسباب الأربعة هو سبب العلاقة التي ندرسها وانما يدلنا فقط كما سبق ان ذكرنا عن شدة أو ضعف هذه العلاقة .

وقد يكون الغرض من حساب معامل الارتباط بجانب قياس درجة العلاقة بين ظاهرتين الاهتداء الى قيمة الظاهرة التابعة لو عرفنا قيمة الظاهرة المتبوعة . فلو عرفنا مثلا ان هناك علاقة قوية بين سن العامل وأجرة لأمكننا ان نحدد الأجر الذي يجب ان يعطى لعامل في سن معين ، كما انه لو عرفنا ان هناك علاقه قوية بين الطول والوزن لأمكننا ان نحكم على ما يجب ان يكون عليه وزن شخص معين ذى طول محدد . وهذه العملية — استنباط قيمة متغير اذا عرفنا قيمة متغير آخر له علاقة به أساسها ما يسمى في الاحصاء بخطوط الانحدار وسوف نتعرض لهذا الموضوع فيما بعد . والمهم ان نعرف الآن ان استنباط قيمة ظاهرة اما اذا عرفنا قيمة ظاهرة أخرى لها علاقة بها يكون أمراً نظرياً فقط قد يتحقق فعلا وقد لا يتحقق في الحياة العملية لأننا ذكرنا سابقاً ان الارتباط يعني **ميل** ظاهرة للتغير اذا تغيرت ظاهرة أخرى .

والارتباط اما ان يكون مستقيماً ويقاس بمعامل الارتباط واما ان يكون

غير مستقيم ويقاس بدليل الارتباط ونسبة الارتباط، فالأول هو ما لم يختلف نوعه من أول المدى الى آخره في الظاهرتين ، والثاني هو ما اختلف فيه طبيعة الارتباط خلال المدى فالأرتباط بين مدة الخدمة في الحكومة والماهية التي يتقاضاها الموظف ارتباط من نوع واحد من أول الى آخر التغير في الظاهرتين فهو لذلك ارتباط مستقيم . أما ارتباط بين كمية ما ينتجه عامل في مصنع في وحدة الزمن وبين طول وقت العمل ارتباط غير مستقيم فهو في بادىء الأمر ارتباط طردي حيث ان كمية المنتج تزيد كلما مكث العامل في عمله وقتاً أطول ثم تأتي اللحظة التي يصبح بعدها الارتباط سالباً حيث تنقص كمية المنتج كلما زاد وقت العمل .

وهناك تقسيم آخر للأرتباط تبعاً لطبيعة المتغيرات نفسها فهناك الأرتباط العادي وهو ما كان بين ظواهر يمكن التعبير عنها رقمياً كما هو الحال في مسائل الأجر والسن وعدد الأولاد ، وما الى ذلك من المتغيرات الرقمية . وهناك الأرتباط الغير رقمي وهو ما كان المقصود منه دراسة العلاقة بين الظواهر غير الرقمية كتحديد العلاقة بين درجة التعليم ودرجة الاهتمام بالرياضة مثلاً وغير ذلك من الأمثله وفي هذه الحالات يسمى مقياس العلاقة بمعامل الأقتران ومعامل التوافق وسيأتي الكلام عليها فيما بعد .

وعند بحث الارتباط بين ظاهرة وأخرى نجد في كثير من الأحوال ظواهر أخرى ذات علاقات أخرى هامة بها وتؤثر على كل منها وعلى علاقتهما ببعضهما فلا يمكن صرف النظر عنها، وعند ئذ اما ان نحسب معامل الأرتباط المتعدد للظاهرة التي ندرسها وحدها مع الظواهر الأخرى المؤثرة عليها مجتمعة . أو ان نحسب معامل الأرتباط الجزئي بين ظاهرة واحدة وظاهرة أخرى أو أكثر مع حذف تأثير باقي الظواهر وذلك طبعاً حسب حاجة البحث .

ومعامل الأرتباط الذي يقيس لنا العلاقة بين الظواهر المختلفة قد صيغ جبريا بحيث لا تزيد قيمته عن الواحد الصحيح،فكلما قرب من الواحد الصحيح

معامل الارتباط جبرياً لا يزيد عن الواحد الصحيح وهو طردي أو عكسي حسب الاشارة الجبرية

كلما دل ذلك على قوة الارتباط . كما انه سمح لهذا المعامل بأن يأخذ اشارة جبرية موجبة أو سالبة للدلاله على طبيعة الأرتباط من حيث انه طردي أو عكسي .

معامل بيرسون للارتباط :

نفترض أن لدينا ظاهرتان س ، ص و ان كل منهما تتخذ القيم المتقابلة الآتية :

الظاهرة	القيمة
$س_1$	$ص_1$
$س_2$	$ص_2$
$س_3$	$ص_3$
$س_4$	$ص_4$
$س_ن$	$ص_ن$

وحيث أن مفهوم الارتباط هو مقارنة التغير الذي يحدث في الظاهرتين فاننا نقيس هذا التغير بحساب انحراف كل قيمة لكل من الظاهرتين عن وسطها الحسابي ، فاذا افترضنا أن الوسط الحسابي للظاهرة س هو $\bar{س}$ وان الوسط الحسابي للظاهرة ص هو $\bar{ص}$ يكون التغير المتقابل في الظاهرتين هو كالآتي :

$س_1 - \bar{س}$	$ص_1 - \bar{ص}$
$س_2 - \bar{س}$	$ض_2 - \bar{ص}$
$س_3 - \bar{س}$	$ص_3 - \bar{ص}$
$س_4 - \bar{س}$	$ص_4 - \bar{ص}$
. . .	. . .
$س_ن - \bar{س}$	$ص_ن - \bar{ص}$

وحيث أن كل قيمة من هذه التغيرات هي قيمة مطلقة فلا يمكن بذلك مقارنة تغيرات الظاهرة س بتغيرات الظاهرة ص ولذلك نحول هذه التغيرات الى تغيرات نسبية بقسمة كل انحراف على الانحراف المعياري للظاهرة الخاصة به كالآتي :

$$\frac{س_1 - \bar{س}}{ع_س} \qquad \frac{ص_1 - \bar{ص}}{ع_ص}$$

$$\frac{س_2 - \bar{س}}{ع_س} \qquad \frac{ص_2 - \bar{ص}}{ع_ص}$$

$$\frac{س_3 - \bar{س}}{ع_س} \qquad \frac{ص_3 - \bar{ص}}{ع_ص}$$

$$\frac{س_4 - \bar{س}}{ع_س} \qquad \frac{ص_4 - \bar{ص}}{ع_ص}$$

$$\frac{\frac{س - \bar{س}}{ن}}{ع_س} \qquad \frac{\frac{ص - \bar{ص}}{ن}}{ع_ص}$$

تقوم فكرة معامل بيرسون للارتباط على أساس انه اذا كان كل تغير في الظاهرة س يصحبه تغير في الظاهرة ص فان ضرب هذه الانحرافات يزيد من الناتج تطبيقاً لفكرة المغناطيسة . كذلك اذا كان كل تغير موجب في س يصحبه تغير موجب في ص فان ناتج الضرب يكون موجباً ويدل ذلك على ارتباط طردي - كذلك إذا كان كل تغير سالب في

س يصحبه تغير سالب في ص فان ناتج الضرب يكون موجباً ويدل ذلك أيضاً على ارتباط طردي . أما إذا كان التغير الموجب في س يصحبه تغير سالب في ص أو العكس فان ناتـج الضرب يكون سالب ويدل ذلك على ارتباط عكسي . وبذلك يرى بيرسون ان ضرب انحرافات س عن وسطها الحسابي في انحرافات ص عن وسطها الحسابي يمكن أن يدلنا على قوة الارتباط وعلى نوعه إذا كان طردياً أو عكسياً (المقياس يسمي معامل الارتباط وأقصى حد له هو واحد صحيح سواء بالزائد أو بالناقص ويدل على وجود ارتباط كامل وكلما بعدت النتيجة عن واحد صحيح كلما قل الارتباط) . وحيث اننا في دراسة الارتباط لا يكون اهتمامنا موجهاً إلى كل حالة على حدة وانما إلى المتوسط العام لحاصل ضرب الانحرافات لذلك نحسب متوسط حصل ضرب انحرافات س في انحرافات ص كالآتي :

$$\frac{1}{\text{ن}}\left(\frac{\text{س}_1-\bar{\text{س}}}{\text{ع}_{\text{س}}}\times\frac{\text{ص}_1-\bar{\text{ص}}}{\text{ع}_{\text{ص}}}+\frac{\text{س}_2-\bar{\text{س}}}{\text{ع}_{\text{س}}}\times\frac{\text{ص}_2-\bar{\text{ص}}}{\text{ع}_{\text{ص}}}\ldots\frac{\text{س}_{\text{ن}}-\bar{\text{س}}}{\text{ع}_{\text{س}}}\times\frac{\text{ص}_{\text{ن}}-\bar{\text{ص}}}{\text{ع}_{\text{ص}}}\right)$$

وبذلك يكون معامل الارتباط $= \dfrac{\text{مج}\,(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}})}{\text{ن}\,\text{ع}_{\text{س}}\,\text{ع}_{\text{ص}}}$ (١)

ومن الواضح انه يمكن وضع هذا القانون في شكل آخر كالآتي :

$$\text{ر} = \frac{\frac{1}{\text{ن}}\ \text{مج}\ (\text{س} - \bar{\text{س}})\ (\text{ص} - \bar{\text{ص}})}{\text{ع}_{\text{س}}\ \ \text{ع}_{\text{ص}}} \qquad (٢)$$

وبالمعالجة الجبرية للقيمة $\text{مج}\ (\text{س} - \bar{\text{س}})\ (\text{ص} - \bar{\text{ص}})$ نحصل على الآتي :

حاصل ضرب انحرافات قيمة س وقيمة ص الاولى =

$$\text{س}_1 \text{ص}_1 - \text{س}_1 \bar{\text{ص}} - \text{ص}_1 \bar{\text{س}} + \bar{\text{س}}\, \bar{\text{ص}}$$

حاصل ضرب انحرافات قيمة س وقيمة ص الثانية =

$$\text{س}_2 \text{ص}_2 - \text{س}_2 \bar{\text{ص}} - \text{ص}_2 \bar{\text{س}} + \bar{\text{س}}\, \bar{\text{ص}}$$

حاصل ضرب انحرافات قيمة س وقيمة ص الثالثة =

$$\text{س}_3 \text{ص}_3 - \text{س}_3 \bar{\text{ص}} - \text{ص}_3 \bar{\text{س}} + \bar{\text{س}}\, \bar{\text{ص}}$$

وبالجمع نحصل على $\text{مج س ص} - \bar{\text{ص}}\ \text{مج س} - \bar{\text{س}}\ \text{مج ص} + \text{ن}\ \bar{\text{س}}\, \bar{\text{ص}}$

وبالتعويض عن مج س ومج ص نحصل على :

$$= \text{مج س ص} - \text{ن}\ \bar{\text{س}}\, \bar{\text{ص}} - \text{ن}\ \bar{\text{س}}\, \bar{\text{ص}} + \text{ن}\ \bar{\text{س}}\, \bar{\text{ص}}$$

$$= \text{مج س ص} - \text{ن}\ \bar{\text{س}}\, \bar{\text{ص}} \qquad (٣)$$

$$\therefore \text{ر} = \frac{\text{مج س ص} - \text{ن}\ \bar{\text{س}}\, \bar{\text{ص}}}{\text{ن}\ \text{ع}_{\text{س}}\ \ \text{ع}_{\text{ص}}}$$

ويمكن وضع هذا القانون في صيفة أخرى كالآتي :

$$ر = \frac{\frac{\text{مجـ س ص}}{ن} - \bar{س}\ \bar{ص}}{ع_{س}\ ع_{ص}} \qquad (٤)$$

واذا أخذنا انحرافات س عن وسط فرضي و وانحرافات ص عن وسط فرضي وَ نحصل على القانون .

$$= \frac{\text{مجـ}\ (س - و)\ (ص - وَ) - ن ح_{س}\ ح_{ص}}{ن\ ع_{س}\ ع_{ص}}$$

حيث $ح_{س} = \bar{س} - و$

$ح_{ص} = \bar{ص} - وَ$

ويمكننا أن نثبت أن هذا القانون لا يختلف عن القوانين السابقة حيث انه ليس الا اشتقاق جبري منها وذلك باجراء العمليات الجبرية الآتية :

حاصل ضرب انحرافات قيمة س وقيمة ص الاولى $= س_{١}\ ص_{١} - ص_{١}\ و - س_{١}\ وَ + و وَ$

» » » » » » الثانية $= س_{٢}\ ص_{٢} - ص_{٢}\ و - س_{٢}\ وَ + و وَ$

» » » » » » الثالثة $= س_{٣}\ ص_{٣} - ص_{٣}\ و - س_{٣}\ وَ + و وَ$

وبالجمع نحصل على مجـ س ص − و مجـ ص − وَ مجـ س + ن و وَ

وبالتعويض عن مجـ س ، مجـ ص نحصل

$$\text{مجـ س ص} - ن\ و\ \bar{ص} - ن\ وَ\ \bar{س} + ن\ و\ وَ$$

والآن ننقل الى الجزء الثاني من القانون $ن ح_{س}\ ح_{ص}$ وهو يساوي

$$ن\ (\bar{س} - و)\ (\bar{ص} - وَ)$$

$$= ن\ \bar{س}\ \bar{ص} - ن\ \bar{س}\ وَ - ن\ و\ \bar{ص} + ن\ و\ وَ$$

بطرح هذا الجزء من الجزء السابق طبقاً للقانون وبذلك نحصل على :

مج س ص ـ ن و صـ ـ ن وـ سـ + ن و وـ ـ ن سـ صـ +

ن سـ وـ + ن و صـ ـ ن و وـ

وهو يساوي مج س ص ـ ن سـ صـ وهو بسط القانون . رقم ٤ .

وبذلك نلاحظ ان هذه القوانين ليست إلا صورا مختلفة من القانون الاول ، إلا ان كلاً منها يمكن أن يسهل علينا العمليات الحسابية التي تتفق مع أنواع البيانات الاحصائية التي تعرض علينا وسوف نرى ذلك في الامثلة التالية ، فإذا كانت انحرافات القيم عن أوساطها الحسابية ارقاماً صغيرة يكون استخدام القانون الاول أمراً بسيطاً من ناحية العمليات الحسابية التي يحتاجها ، أما إذا كانت القيم صغيرة ولكن انحرافاتها عن أوساطها الحسابية انحرافات كبيرة (رقمياً) يكون استخدام القانون الثاني أسهل ، أما اذا كانت القيم كبيرة وانحرافاتها عن أوساطها الحسابية قيما كبيرة (رقمياً) يكون استخدام القانون الثالث أسهل . وفي التوزيعات التكرارية تكون العمليات الحسابية معقدة وطويلة إذا استخدمنا القانونين الاول والثاني ولذلك نقتصر دائماً في هذه الحالة على استخدام القانون الثالث لأنه يبسط العمل الحسابي كثيراً . ولذلك تكون المفاضلة بين هذه القوانين هي فقط في حالة القيم غير المبوبة ، أما في حالة القيم المبوبة يجب أن يقتصر تفكيرنا على القانون الثالث. والمفاضلة هي فقط في ناحية سهولة أو صعوبة العمليات الحسابية ، إذ اننا نصل الى نفس النتيجة لمعامل الارتباط مهما كان القانون الذي نستخدمه من هذه القوانين الثلاث ، فهي كما قدمنا ليست إلا صوراً جبرية مختلفة لشيء واحد .

قانون سبيرمان للارتباط :

وضع سبيرمان قانوناً للارتباط على أساس ترتيب قيم كل من الظاهرتين س ، ص المتقابلتين في الشكل $ر = \frac{١ - ٦ \text{ مج } ف^{٢}}{ن (ن^{٢} - ١)}$ وقد استنبطه من

القانون $ر = \frac{\text{مج س ص} - \text{ن } \bar{\text{س}}\,\bar{\text{ص}}}{\text{ن }ع_{\text{س}}\,ع_{\text{ص}}}$ كالآتي :

نفترض ان ترتيب قيم س هي ، ١ ، ٢ ، ٣ ، ٤ .... ن

وان ترتيب قيم ص المناظرة هي ٣ ، ٤ ، ١ ، ٢ ... . . ن وهي نفس الأرقام السابقة الخاصة بالظاهرة س ولكن في وضع مختلف .

$$\text{الوسط الحسابي لقيم س} = \frac{١ + ٢ + ٣ + \cdots + \text{ن}}{\text{ن}} =$$

$$\frac{\frac{\text{ن}(\text{ن} + ١)}{٢}}{\text{ن}} \quad \text{(نظرية جبرية)}$$

$$= \frac{\text{ن}(\text{ن} + ١)}{٢\,\text{ن}} = \frac{\text{ن} + ١}{٢}$$

وحيث ان قيم ص هي نفس القيم وان كانت بوضع مختلف يكون الوسط الحسابي لها كذلك :

$$\frac{\text{ن} + ١}{١}$$

أي ان $\bar{\text{س}} = \bar{\text{ص}}$

$$ع^{٢}_{\text{س}} = \frac{\text{مج س}^{٢}}{\text{ن}} - \bar{\text{س}}^{٢}$$

$$\text{مجـ س}^{٢} = ١^{٢} + ٢^{٢} + ٣^{٢} + \cdots + \text{ن}^{٢} =$$

$$\frac{\text{ن}\,(\text{ن} + ١)\,(٢\,\text{ن} + ١)}{٦\,\text{ن}} \quad (\text{نظرية جبرية})$$

$$\therefore \text{ع}^{٢}\text{س} = \frac{\text{ن}\,(\text{ن} + ١)\,(٢\,\text{ن} + ١)}{٦\,\text{ن}} - \left(\frac{\text{ن} + ١}{٢}\right)^{٢}$$

$$= \frac{(\text{ن} + ١)\,(٢\,\text{ن} + ١)}{٦} - \frac{\text{ن}^{٢} + ٢\text{ن} + ١}{٤}$$

$$= \frac{٤\,\text{ن}^{٢} + ٦\,\text{ن} + ٢ - ٣\,\text{ن}^{٢} - ٦\,\text{ن} - ٣}{١٢}$$

$$= \frac{\text{ن} - ١}{١٢}$$

$$\text{وكذلك ع}^{٢}\,\text{ص} = \frac{\text{ن}^{٢} - ١}{١٢}$$

$$\text{، مجـ ف}^{٢} = \text{مجـ}\,(\text{س} - \text{ص})^{٢}$$

$$= \text{مجـ س}^{٢} - ٢\,\text{مجـ س ص} + \text{مجـ ص}^{٢}$$

$$\text{وحيث أن مجـ س}^{٢} = \text{مجـ ص}^{٢}\ \text{وكل منهما} = \frac{\text{ن}\,(\text{ن} + ١)\,(٢\text{ن} + ١)}{٦}$$

$$\therefore \text{مجـ ف}^{٢} = \frac{٢\text{ن}\,(\text{ن} + ١)\,(٢\,\text{ن} + ١)}{٦} - ٢\,\text{مجـ س ص}\ .$$

$$\therefore \text{مجـ س ص} = \frac{\text{ن}\,(\text{ن} + ١)\,(٢\text{ن} + ١)}{٦} - \frac{١}{٢}\,\text{مجـ ف}^{٢}$$

وحيث أن $\text{ر} = \dfrac{\text{مج س ص} - \text{ن } \bar{\text{س}}\,\bar{\text{ص}}}{\text{ن } \text{ع}_{\text{س}}\,\text{ع}_{\text{ص}}}$

∴ بالتعويض عن مج س ص ، $\bar{\text{س}}$ ، $\bar{\text{ص}}$ ، $\text{ع}_{\text{س}}$ ، $\text{ع}_{\text{ص}}$ نحصل على :

$$\text{ر} = \frac{\dfrac{\text{ن}(\text{ن}+١)(٢\text{ن}+١)}{٦} - \dfrac{١}{٢}\,\text{مج ف}^{٢} - \text{ن}\left(\dfrac{\text{ن}+١}{٢}\right)^{٢}}{\text{ن}\sqrt{\dfrac{(\text{ن}^{٢}-١)}{١٢} \times \dfrac{(\text{ن}^{٢}-١)}{١٢}}}$$

$$\therefore \text{ر} = \frac{\dfrac{٢\text{ن}^{٣} - ٢\text{ن}^{٢} + \text{ن}}{٦} - \dfrac{\text{ن}^{٣} + ٢\text{ن}^{٢} + \text{ن}}{٤} \quad \dfrac{١}{٢}\,\text{مج ف}^{٢}}{\text{ن} \times \dfrac{\text{ن}^{٢} - ١}{١٢}}$$

$$\therefore \text{ر} = \frac{\dfrac{\text{ن}^{٣} - \text{ن}}{١٢}}{\dfrac{\text{ن}^{٣} - \text{ن}}{١٢}} - \frac{\dfrac{١}{٢}\,\text{مج ف}^{٢}}{\dfrac{\text{ن}^{٣} - \text{ن}}{١٢}}$$

$$\therefore \text{ر} = ١ - \frac{٦\ \text{مج ف}^{٢}}{\text{ن}(\text{ن}^{٢} - ١)}$$

تطبيقات على قوانين الارتباط .

مثـــال ٥٣ :

من البيانات الآتية احسب الارتباط :

س	٦٥	٦٣	٦٧	٦٤	٦٨	٦٢	٧٠	٦٦	٦٨	٦٧	٦٩	٧١
ص	٦٨	٦٦	٦٨	٦٥	٦٩	٦٦	٦٨	٦٥	٧١	٦٧	٦٨	٧٠

إذا أردنا تطبيق القانون الاول $ر = \frac{مج (س - \bar{س}) (ص - \bar{ص})}{ن \; ع_س \; ع_ص}$

تكون خطوات العمل كالآتي :

ص	(س - س̄)	(ص - ص̄)	(س - س̄) (ص - ص̄)	(س - س̄)^2	(ص - ص̄)^2
٦٨	− ١ر٧	٠ر٤	− ٠ر٦٨	٢ر٨٩	٠ر١٦
٦٦	− ٣ر٧	− ١ر٦	٥ر٩٢	١٣ر١٩	٢ر٥٦
٦٨	+ ٠ر٣	+ ٠ر٤	ر١٢	ر٠٩	٠ر١٦
٦٥	− ٢ر٧	− ٢ر٦	٧ر٠٢	٧ر٢٩	٦ر٧٦
٦٩	+ ١ر٣	+ ١ر٤	١ر٨٢	١ر٦٩	٥ر٩٦
٦٦	− ٤ر٧	− ١ر٦	٧ر٥٣	٢٢ر٠٩	٢ر٥٦
٦٨	+ ٣ر٣	+ ر٤	١ر٣٢	١٠ر٨٩	٠ر١٦
٦٥	− ٠ر٧	− ٢ر٦	١ر٨٢	ر٤٩	٦ر٧٦
٧١	١ر٣	+ ٣ر٤	٤ر٤٢	١ر٦٩	١١ر٥٦
٦٧	− ٠ر٣	− ٠ر٦	− ٠ر١٨	ر٠٩	٠ر٣٦
٦٨	+ ٢ر٣	+ ٠ر٤	٠ر٩٢	٥ر٢٩	٠ر١٦
٧٠	+ ٤ر٣	+ ٢ر٤	١٠ر٣٢	١٨ر٤٩	٥ر٧٦
٨١١	−	−	٤٠ر٣٤	٨٤ر٦٨	٣٨ر٩٢

$\bar{س} = \frac{٨٠٠}{١٢} = ٦٦ر٧$ $\qquad \bar{ص} = \frac{٨١١}{١٢} = ٦٧ر٦$

$$ر = \frac{٤٠ر٣٤}{١٢\sqrt{\frac{٨٤ر٦٨}{١٢} \times \frac{٣٨ر٩٢}{١٢}}} = \frac{٤٠ر٣٤}{٥٧ر٤٠}$$

∴ ر = ٧ر٠

وإذا أردنا تطبيق القانون الثاني ر = $\frac{\text{مجـ س ص} - \text{ن } \bar{س}\ \bar{ص}}{\text{ن } ع_{س}\ ع_{ص}}$

تكون خطوات العمل كالآتي :

س	ص	س ص	$س^٢$	$ص^٢$
٦٥	٦٨	٤٤٢٠	٤٢٢٥	٤٦٢٤
٦٣	٦٦	٤١٥٨	٣٩٦٩	٤٣٥٦
٦٧	٦٨	٤٥٥٦	٤٤٨٩	٤٦٢٤
٦٤	٦٥	٤١٦٠	٤٠٩٦	٤٢٢٥
٦٨	٦٩	٤٦٩٢	٤٦٢٤	٤٧٦١
٦٢	٦٦	٤٠٩٢	٣٨٤٤	٤٣٥٦
٧٠	٦٨	٤٧٦٠	٣٩٠٠	٤٦٢٤
٦٦	٦٥	٤٢٩٠	٤٣٥٦	٣٢٢٥
٦٨	٧١	٤٨٢٨	٤٦٢٤	٥٠٤١
٦٧	٦٧	٤٤٨٩	٤٤٨٩	٤٤٨٩
٦٩	٦٨	٤٦٩٢	٤٧٦١	٤٦٢٤
٧١	٨٠	٤٩٧٠	٥٠٤١	٤٩٠٠
٨٠٠	٨١١	٥٤١٠٧	٥٣٤١٨	٥٤٨٤٩

$$\therefore \text{ر} = \frac{٥٤١٠٧ - ١٢ \times \frac{٨٠٠}{١٢} \times \frac{٨١١}{١٢}}{١٢\sqrt{\left(\frac{٥٣٤١٨}{١٢} - \left(\frac{٨٠٠}{١٢}\right)^{٢}\right) \times \left(\frac{٥٤٨٤٩}{١٢} - \left(\frac{٨١١}{١٢}\right)^{٢}\right)}}$$

$$\therefore \text{ر} = \frac{٤٠\text{ر}٣٤}{٥٧\text{ر}٤٠} = ٠\text{ر}٧$$

وإذا أردنا تطبيق القـــانون الثالث ر =

$$\frac{\text{مجـ}(\text{س} - \text{و})(\text{ص} - \text{وَ}) - \text{ن}\ \text{ح}_{\text{س}}\ \text{ح}_{\text{ص}}}{\text{ن}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}}}$$

تكون خطوات العمل كالآتي :

س	ص	س − ٦٨	ص − ٦٨	(س − ٦٨) (ص − ٦٨)	(س − ٦٨)^{٢}	(ص − ٦٨)^{٢}
٦٨	٦٨	صفر	صفر	صفر	صفر	صفر
٦٣	٦٦	− ٥	− ٢	+ ١٠	٣٥	٤
٦٧	٦٨	− ١	صفر	صفر	١	صفر
٦٤	٦٥	− ٤	− ٣	+ ١٢	١٦	٩
٦٨	٦٩	صفر	+ ١	صفر	صفر	١
٦٢	٦٦	− ٦	− ٢	+ ١٢	٣٦	٤
٧٠	٦٨	+ ٢	صفر	صفر	٤	صفر
٦٦	٦٥	− ٢	− ٣	+ ٦	٤	٩
٦٨	٧١	صفر	+ ٣	صفر	صفر	٩
٦٧	٦٧	− ١	− ١	+ ١	١	١
٦٩	٦٨	+ ١	صفر	صفر	١	صفر
٧١	٧٠	+ ٣	+ ٣	+ ٦	١	٤
		−	−	٤٧	٩٧	٤١

$$\text{ح}_{\text{س}} = \text{٦٦٫٦٧} - \text{٦٨} = -\ \text{١٫٣٣}$$

$$\text{ح}_{\text{ص}} = \text{٦٧٫٥٨} - \text{٦٨} = -\ \text{٠٫٤٢}$$

$$\therefore\ \text{ر} = \frac{\text{٤٧} - \text{١٢} \times -\ \text{١٫٣٣} \times -\ \text{٠٫٤٢}}{\text{١٢}\ \sqrt{\frac{\text{٩٧}}{\text{١٢}} - (\text{١٫٣٣})^{\text{٢}}}\ \sqrt{\frac{\text{٥١}}{\text{١٢}} - (\text{٠٫٤٢})^{\text{٢}}}}$$

$$= \frac{٤٧ - ٦٫٧}{٥٧٫٤} = \frac{٤٠٫٣}{٥٧٫٤} = ٠٫٧$$

وإذا أردنا تطبيق قانون سبيرمان $١ - \frac{٦ \text{ مج ف}^{٢}}{\text{ن}(\text{ن}^{٢} - ١)}$ تكون خطوات العمل كالآتي :

(نلاحظ أن قيم كل من س ، ص ترتب ترتيباً تنازلياً وبذلك تكون ف هي الفرق بين هذه التراتيب .

س	ص	تراتيب س	تراتيب ص	ف	ف^{٢}
٦٨	٦٨	٥	٥٫٥	٠٫٥	٠٫٢٥
٦٣	٦٦	١١	٩٫٥	١٫٥	٢٫٢٥
٦٧	٦٨	٧٫٥	٥٫٥	٢	٤
٦٤	٦٥	١٠	١١٫٥	١٫٥	٢٫٢٥
٦٨	٦٩	٥	٣	٢	٤
٦٢	٦٦	١٢	٩٫٥	٢٫٥	٦٫٢٥
٧٠	٦٨	٢	٩٫٥	٣٫٥	١٢٫٢٥
٦٦	٦٥	٩	١١٫٥	٢٫٥	٦٫٢٥
٦٨	٧١	٥	١	٤	١٦
٦٧	٦٧	٧٫٥	٨	٠٫٥	٠٫٢٥
٦٩	٦٨	٣	٥٫٥	٢٫٥	٦٫٢٥
٧١	٧٠	١	٢	١	١
–	–	–	–	–	٦١

(ملاحظة – اذا كانت القيم متشابهة تعطي متوسط التراتيت المسلسلة لها)

$$\therefore \text{ر} = ١ - \frac{٦ \ \text{مج ف}^{٢}}{\text{ن}(\text{ن}^{٢} - ١)}$$

$$= ١ - \frac{٦ \times ٦١}{١٢ \ (١٤٤ - ١)} = ١ - \frac{٣٦٦}{١٧١٦}$$

$$= ١ - ٢١\text{ر} = ٧٩\text{ر} \ ٠$$

واذا كانت القيم مبوبة في توزيع تكراري مزدوج يكون من الافضل ، كما قدمنا أن نقتصر على استخدام القانون الثالث .

$$\text{ر} = \frac{\text{مج} \ (\text{س} - \text{و}) \ (\text{ص} - \bar{\text{و}}) - \text{ن} \ \text{ح}_{\text{س}} \ \text{ح}_{\text{ص}}}{\text{ن} \ \text{ع}_{\text{س}} \ \text{ع}_{\text{ص}}}$$

وتكون خطوات العمل كالآتي :

مثال ٥٤ :

س	ص					المجموع
	٣ ـ	٥ ـ	٧ ـ	٩ ـ	١١ ـ ١٣	
صفر ـ	٩	٢٩				٣٨
١ ـ	١	١١	١٢			٢٤
٢ ـ			١٣	٥		١٨
٣ ـ				١٥	٥	٢٠
المجموع	١٠	٤٠	٢٥	٢٠	٥	١٠٠

التوزيع الخاص بالظاهرة س :

ف	ك	م	ح	ح ك	ح٢ ك
صفر —	٣٨	٠ر٥	— ١	— ٣٨	٣٨
١ —	٢٤	١ر٥	صفر	صفر	صفر
٢ —	١٨	٢ر٥	+ ١	+ ١٨	١٨
٣ - ٤	٢٠	٣ر٥	+ ٢	+ ٤٠	٨٠
المجموع	١٠٠	—	—	٢٠	١٣٦

$$\bar{س} = ١ر٥ + \frac{٢٠}{١٠٠} \times ١ = ١ر٧$$

$$ع_{س} = \sqrt{١ر٣٦ - ٠ر٠٤} = \sqrt{١ر٣٢}$$

التوزيع الخاص بالظاهرة ص :

ف	ك	م	ح	ح ك	ح٢ ك
٣ —	١٠	٤	— ٢	— ٢٠	٤٠
٥ —	٤٠	٦	— ١	— ٤٠	٤٠
٧ —	٢٥	٨	صفر	صفر	صفر
٩ —	٢٠	١٠	+ ١	+ ٢٠	٢٠
١١ — ١٣	٥	١٢	+ ٢	+ ١٠	٢٠
المجموع	١٠٠	—	—	— ٣٠	١٢٠

$$\text{ص}^{-} = ٨ - \frac{٣٠}{١٠٠} \times ٢ = ٧٫٤$$

$$\text{ع}_{\text{ص}} = ٢ \sqrt{١٫٢ - \left(\frac{٣٠}{١٠٠}\right)^{٢}} = ٢ \sqrt{١٫١١}$$

ولحساب مجـ (س − و) (ص − و) نرجع إلى الجـــدول المزدوج ، ومنه نكون الجدول الآتي الذي نضع فيه كل تكرار ورد في جسم الجدول (مـــا عدا المجاميع) وأمامه انحرافه الواقعي الخاص بالظاهرة س وانحرافه الواقعي الخاص بالظاهرة ص ونضرب الارقام الثلاث في بعضها كالآتي :

ك	س − و	ص − و−	ك (س − و) (ص − و−)
٩	− ٤	− ١	٣٧
١	− ٤	صفر	صفر
٢٩	− ٢	− ١	٥٨
١١	− ٢	صفر	صفر
١٢	صفر	صفر	صفر
١٣	صفر	+ ١	صفر
٥	+ ٢	+ ١	١٠
١٥	+ ٢	+ ٢	٦٠
٥	+ ٤	+ ٢	٤٠
١٠٠	—	—	٣٠٤

مـــلاحظة :

يمكن اجراء هذه العملية على الجدول المزدوج نفسه حيث نضع أمـــام

فئات س وانحرافاتها عن وسطها الفرضي (يجب ان تكون انحرافات واقعية غير مختصرة على طول الفئة) ، وأمـــام فئات ص انحرافاتها عن وسطها الفرضي (يجب أن تكون انحرافات واقعية غير مختصرة على طول الفئة) . أمام الانحراف السيني صفر نشطب جميع الأرقام في جسم الجدول وكذلك أمام الانحراف الصادي صفر نشطب جميع الأرقام في جسم الجدول ، حيث ان هذه الأرقام سوف تضرب في صفر فيكون ناتجها = صفر . بعد ذلك نأخذ كل تكرار في جسم الجدول ونضربه في الانحراف السيني المقابل له في الانحراف الصادي المقابل له (لاحظ ان ثلاث أرقام تضرب في بعضها) ونضع النتيجة لدينا باشارتها الجبرية ثم نجمع هذه النتائج جمعاً جبرياً .

على أساس الناتج السابق يكون معامل الارتباط ر كالآتي :

$$ر = \frac{١٠٤ - ١٠٠ \times ر٢ \times - ر٦}{١٠٠ \times ١ \times ٢ \sqrt{١ر١١ \times ١ر٣٢}}$$

لاحظ ان $ح_س = \bar{س} - و$ وحيث ان $\bar{س} = و + \frac{مجـ\ ح\ ك}{مجـ\ ك}$ × طول الفئة .

$\therefore \bar{س} - و = \frac{مجـ\ ح\ ك}{مجـ\ ك}$ × طول الفئــة ، وكذلك بالنسبة للظاهرة

حيث $ح_ص = \bar{ص} - و$ وحيث ان $\bar{ص} = \bar{و} + \frac{مجـ\ ح\ ك}{مجـ\ ك}$ × طول الفئة

$\therefore \bar{ص} - \bar{و} = \frac{مجـ\ ح\ ك}{مجـ\ ك}$ × طول الفئة .

$$\therefore ر = \frac{٢١٦}{٢٤٢ر١} = ٠ر٨٩$$

ويمكن اجراء العمل باستعمال نفس القانون ولكن بشكـل آخر في الجدول الآتي :

س	ص: ـ٣	ـ٥	ـ٧	ـ٩	١١ ـ ١٣	$ك_س$	$ح_س$	$ح_س ك_س$	$ح^٢_س ك_س$	مجـ $ح_ص$	مجـ $ح_س ح_ص$
صفر ـ	٩	٢٩				٣٨	ـ ١	ـ ٣٨	٣٨	ـ ٩٤	٩٤
١ ـ	١	١١	١٢			٢٤	صفر	صفر		ـ ٢٣	صفر
٣ ـ			١٣	٥		١٨	+ ١	+ ١٨	١٨	+ ١٠	١٠
٣ ـ ٤				١٥	٥	٢٠	+ ٢	+ ٤٠	٨٠	+ ٥٠	١٠٠
$ك_ص$	١٠	٤٠	٢٥	٢٠	٥	١٠٠	ـ	+ ٢٠	١٣٦	ـ ٦٠	٢٠٤
$ح_ص$	ـ ٤	ـ ٢	صفر	+ ٢	+ ٤	ـ					
$ح_ص ك_ص$	ـ ٤٠	ـ ٨٠	صفر	+ ٤٠	+ ٢٠	ـ ٦٠					
$ح^٢_ص ك_ص$	ـ ١٦	ـ ١٦	صفر	ـ ٨٠	٨٠	ـ ٤٨					
مجـ $ح_س$	ـ ٩	ـ ٢٩	+ ١٣	+ ٣٥	ـ ١	+ ٢٠					
مجـ $ح_س ح_ص$	+ ٣٦	+ ٥٨	صفر	+ ٧٠	٤٠	٢٠٤					

$ح_س = \frac{٢٠}{١٠٠} = ٢ر٠$

$ع_س = \sqrt{١ر٣٦ - ٠٤ر٠} = \sqrt{١ر٣٢}$

$ح_ص = \frac{-٦٠}{١٠٠} = -٦ر٠$

$ع_ص = \sqrt{٤ر٨ - ٠ر٣٦} = \sqrt{٤ر٤٤}$

$ر = \frac{٢٠٤ - ١٠٠ \times ٢ر٠ \times -٦ر٠}{١٠٠ \sqrt{١ر٣٢ \times ٤ر٤٤}}$

$= \frac{٢١٦}{٢٤٢ر١}$

$= ٨٩ر٠$

معامل ارتباط الرتب للبيانات الوصفية

تصور ان احصائياً اجتماعياً قام بدراسة حالة سبع عائلات مختلفة في حي معين وسجل لكل عائلة الحالة العلميـة لرب الأسرة والمستوى الاقتصادي للأسرة نفسها فخرج بالجدول الآتي :

الحالة العلمية لرب الاسرة	المستوى الاقتصادي للاسرة
يحمل شهادات متوسطة	فقيرة
أمـي	معدمـة
يقرأ أو يكتب	فقيرة
يحمل شهادات عالية	غنيـة
أمـي	معدمـة
أمـي	متوسط الحال
يقرأ ويكتب	فقيرة

فاذا شاء هذا الاحصائي أن يقف على مدى العلاقة بين هاتين الظاهرتين يجب أن يستخدم معامل ارتباط الرتب وهو ما يسمى أحياناً بمعامـل سبيرمان نسبة الى صاحبه حيث ان البيانات المتوفرة لدينا بيانات وصفيـة ويمكن ترتيبها (لا يمكن استخدام معامل بيرسون حيث ان هذا المعامل لا يمكن تطبيقه الا في حالة البيانات القيمية) .

والفكرة الأساسية في قياس معامل ارتباط الرتب ، كما قدمنـا ، هي مقارنة رتبتي الأسرة الواحدة في الظاهرتين فان اختلفتا كثيراً دل ذلك على قلة الارتباط وان اتحدا دل ذلك على شدة الارتباط أي ان أساس المعامل هو الفروق بين الرتب المتقابلة فكلما كبرت هذه الفروق في المتوسط كلما ضعف الارتباط بين الظاهرتين والعكس كلما صغرت هذه الفروق .

ومعامل سبيرمان الذي يقوم على هذا الأساس هو $١ - \frac{٦ \text{ مجـ } ف^{٢}}{ن(ن^{٢} - ١)}$

حيث $ف^{٢}$ = مجموع مربعات الفروق بين الرتب المتقابلة

ن = عدد وحدات البحث .

ولحساب هذه المعامل نرتب الظاهرتين ترتيباً تدريجياً منتظما ، وفي مثالنا السابق نعطي كل أسرة رتبـــة حسب الحالة العلمية لرب الاسرة والحالة الاقتصادية للأسرة نفسها ثم نحسب الفرق بين الرتبتين ثم نربع هذا الفرق لزيادة حساسية المعامل ثم نطبق القاعدة السابقة ـ ويكون العمل كالآتي :

رتبة الحالة العملية		رتبة الحالة الاقتصادية		الفرق بين الرتب المعدلة	$ف^{٢}$
أولية	معدلة	أولية	معدلة	ف	
٦	٦	(١)	٤	٢	٤ر٠٠
(١)	٢	(١)	١ر٥	٠ر٥	٠ر٢٥
(٤)	١ر٥	(٤)	٤	٠ر٥	٠ر٢٥
٧	٧	٧	٧	٠	٠٠
(٢)	٢	(٢)	١ر٥	٠ر٥	٠ر٢٥
(٣)	٢	٦	٦	٤	١٦ر٠٠
(٥)	٤ر٥	(٥)	٤	٠ر٥	٠ر٢٥
—	—	—	—	—	٢١

ر (معامل ارتباط الرتب) = $١ - \frac{٦ \times ٢١}{٧ (٤٩ - ١)} = ١ - \frac{١٢٦}{٣٣٦}$

= ١ − ٣٧٥ر = ٦٢٥ر

(أي ان الارتباط طردي قوي نوعاً) .

نلاحظ ان الرتب المتشابهة وضعناها بين أقواس حسب ترتيبها الجاري في المجموعة ثم عدلنا هذه الرتب بأخذ المتوسط لها ، فالحالة العلمية يقرأ ويكتب تكررت مرتين وترتيبها هو ٤ ، ٥ ، فأخذنا المتوسط ٤ر٥ كترتيب معدل . والحالة أمي تكررت ثلاث مرات وترتيبها ١ ، ٢ ، ٣ فأخذنا المتوسط ٢ كترتيب معدل وهكذا بالنسبة للظاهرة الثانية وهي الخاصة بالحالة الاقتصادية للأسرة نفسها .

وميزة هذا المعامل هي بساطة حسابه ، فالعمليات الحسابية التي يحتاجها سهلة للغاية ولكن لا يجب أن نستخدمه في قياس الارتباط الا في حـــالة ما يكون ترتيب الظواهر أمراً منطقياً معقولا .

معامل الارتباط للقيم المبوبة بطريقة الاقطار ذات الفروق المتساوية او المجاميـع المتساوية :

يمكن ايجاد معامل الارتباط من بيانات مبوبة بطريقة أسهل من طريقة بيرسون وتؤدي إلى نفس النتيجة تماماً ، وذلك بأن نرتب فئات س تصاعديا أو تنازلياً ونرتب فئات ص بنفس الكيفية ثم نتبع طريقة أقطار الفروق المتساوية أو طريقة أقطار المجاميع المتساوية . وتمتاز طريقة الرتب بانهـــا تصلح للاستعمال في الجداول التي تكون بها فئة أو أكثر من الفئات المفتوحة ولكنها لا تصلح مطلقاً للاستعمال في جدول تكون به الفئات غير متساوية .

وتعتمد هذه الطريقة على حساب الانحرافات المعيارية لرتب المتغيرين وللفرق بين رتبتيها أو لمجموعها . وتنحصر خطوات العمل في اعطاء كل فئة من فئات المتغيرين رتبة تبدأ بالواحد الصحيح وتنتهي بعدد مساو لعدد فئات المتغير . ثم رسم اقطار في الجدول المزدوج تمر من اليمين إلى اليسار ويمتاز كل قطر منها بتساوي الفرق بـــين الرتبة السينية والصادية في جميـع أجزائه . ويحتاج حساب معامل الارتباط بعد ذلك لتطبيق القانون التالي :

$$ر = \frac{ع^{٢}_{س} + ع^{٢}_{ص} - ع^{٢}_{ف}}{٢\ ع_{س}\ ع_{ص}}$$

حيث $ع^{٢}_{س}$ مربع الانحراف المعياري لرتب س

حيث $ع^{٢}_{ص}$ مربع الانحراف المعياري لرتب ص

حيث $ع^{٢}_{ف}$ مربع الانحراف المعياري للفرق بين رتب س ، ورتب ص
وسيتضح استخدام هـا القانون من المثال الآتي :

مثـــال ٥٦ :

رتب ص	فئات الايجار / فئات الرتب	٣ـ	٥ـ	٧ـ	٩ـ	١١ـ	المجموع
رتب س		١	٢	٣	٤	٥	
١	أقل من ١	٩	٢٩				٣٨
٢	١ ـ	١	١١	١٢			٢٤
٣	٢ ـ			١٣	٥		١٨
٤	٣ ـ				١٥	٥	٢٠
	المجموع	١٠	٤٠	٢٥	٢٠	٥	١٠٠

أولا : حساب $ع^{٢}_{س}$

رتب س	ك	ح	ح ك	ح٢ ك
١	١٠	— ٢	— ٢٠	٤٠
٢	٤٠	— ١	— ٤٠	٤٠
٣	٢٥	صفر	صفر	صفر
٤	٢٠	+ ١	٢٠	٢٠
٥	٥	+ ٢	١٠	٢٠
المجموع	١٠٠	—	— ٦٠ + ٣٠ — ٣٠	

$$ع^{٢}_{س} = \sqrt{\frac{١٢٠}{١٠٠} - \left(\frac{-٣٠}{١٠٠}\right)^{٢}} = \sqrt{١ر٢ - ٠٩ر} = \sqrt{١ر١١}$$

ثانياً : حساب $ع^{٢}_{ص}$

رتب ص	ك	ح	ح ك	ح^{2} ك
١	٣٨	− ٢	− ٧٦	١٥٢
٢	٢٤	− ١	− ٢٤	٢٤
٣	١٨	صفر	صفر	صفر
٤	٢٠	+ ١	٢٠	٢٠
المجموع	١٠٠	−	− ١٠٠	١٩٦
			+ ٢٠	
			− ٨٠	

$$ع^{2}_{ص} = \sqrt{\frac{١٩٦}{١٠٠} - \left(\frac{٨٠}{١٠٠}\right)^{2}} = \sqrt{١٫٣٢}$$

ثالثاً : حساب $ع^{2}_{ف}$

الفرق بين رتبتي س ، ص		ك	ك ف	ك ف^{2}
القطر				
ا	١	٢٩ + ١٢ + ٥ + ٥ = ٥١	٥١	٥١
ب	صفر	٩ + ١١ + ١٣ + ١٥ = ٤٨	صفر	صفر
ج	− ١	١	− ١	١
المجموع		١٠٠	٥٠	٥٢

$$\text{ع}_{\text{٢}}\ \text{ف} = \sqrt{\frac{\text{٥٢}}{\text{١٠٠}} - \left(\frac{\text{٥٠}}{\text{١٠٠}}\right)^{\text{٢}}} = \sqrt{\text{٥٢ر} - \text{٢٥ر}} = \sqrt{\text{٢٧ر}}$$

$$\therefore \text{ر} = \frac{\text{١١ر١} + \text{٣٢ر١} - \text{٢٧ر}}{\text{٢}\sqrt{\text{١١ر١} \times \text{٣٢ر١}}} = \frac{\text{٤٣ر٢} - \text{٢٧ر}}{\sqrt{\text{٤٦٥٢ر١}}}$$

$$= \frac{\text{١٦ر٢}}{\text{٢}\ (\text{٢١٠٤ر١})} = \text{٨٩ر}$$

ويمكن استخدام قانون آخر هو $\text{ر} = \dfrac{\text{ع}^{\text{٢}}_{\text{ح}} - \text{ع}^{\text{٢}}_{\text{س}} - \text{ع}^{\text{٢}}_{\text{ص}}}{\text{٢}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}}}$

حيث $\text{ع}^{\text{٢}}_{\text{س}}$ مربع الانحراف المعياري لرتب س

، $\text{ع}^{\text{٢}}_{\text{ص}}$ » » » » ص

، $\text{ع}^{\text{٢}}_{\text{ح}}$ » » » » لمجموع رتب س،ص

ولتطبيق هذا القانون نرسم أقطاراً تمر من اليسار إلى اليمين ويمتاز كل قطر منها بتساوي مجموع رتب س ، ص في جميع أجزائه .

القطر	مجموع الرتب	ك	ح	ك ح	ك ح٢
أ	٢	٩	−٣	−٢٧	٨١
ب	٣	٣٠	−٢	−٦٠	١٢٠
جـ	٤	١١	−١	−١١	١١
د	٥	١٢	صفر	صفر	صفر
هـ	٦	١٣	+١	١٣	١٢
و	٧	٥	+٢	١٠	٢٠
ل	٨	١٥	+٣	٤٥	١٣٥
م	٩	٥	+٤	٢٠	٨٠
المجموع	−	−	−	−٩٨ +٨٨ −١٠	٤٦٠

$$\text{ع}^{٢}_{\text{ح}} = \sqrt{\frac{\text{٤٦٠}}{\text{١٠٠}} - \left(\frac{\text{١٠٠}}{\text{١٠٠}}\right)^{٢}}$$

$$= \sqrt{\text{٤ر٦} - \text{ر٠١}} - \sqrt{\text{٤ر٥٩}}$$

$$= \frac{\text{٤ر٥٩} - \text{١ر٣٢} - \text{١ر١١}}{\text{٢}\sqrt{\text{١ر١١} \times \text{١ر٣٢}}} = \frac{\text{٢ر١٦}}{\text{٢ (١ر٢١٠٤)}} = \text{ر٨٩}$$

الارتباط بين الظواهر غير الرقمية :

نحتاج أحياناً إلى دراسة درجة العلاقة بين ظاهرتين لا يمكن التعبير عنهما

بالارقام مثل الارتباط بين التطعيم بمصل ضد مرض معين والاصابة بهذا المرض أو بين الحالة التعليمية لمجموعة من الأشخاص ومستواهم المادي وغير ذلك من الأمثلة الكثيرة ، وفي هذه الحالةلا نستطيع أن نستخدم معامل الارتباط أو معامل الائتلاف لأن حساب هـذين المعاملين يتوقف على معرفة الأوساط الحسابية وغير ذلك من المقاييس التي لا يمكن حسابها في حالة الظواهر غير الرقمية .

ولقياس الارتباط بين هذه الظواهر نستخدم معامل الاقتران والقانون الخاص به هو

$$ن = \frac{ا د - ب حـ}{ا د + ب حـ}$$

ولشرح هذا القاتون نأخذ مثالاً عملياً مع ملاحظة ان هذا القانون لا يمكن استخدامه إلا في الحالات التي تنقسم فيها الظواهر التي ندرس العلاقة بينها إلى نوعين فقط ، أما فيما عـدا ذلك من الحالات يجب استخدام مقياس آخر سيأتي الكلام عنه فيما بعد . فـاذا فرضنا انه أثناء وباء للتيفود مثلاً أجرى أحد الأطباء تجربة مصل جديد على عينة من الأفراد عددها ن ، وقد طعم منهم بالمصل عدداً معيناً فاصيب بالمرض العدد حـ ولم يصب العدد ا — أمـا الباقي الذين لم يطعموا فأصيب منهم العدد د ولم يصب العدد ب ، فكيف يمكن من هذه البيانات أن ندرس درجة العلاقة بين الحقن بالمصل وعدم الاصابة .

يمكننا وضع البيانات السابقة في الجدول التالي .

مثـــال ٥٧ :

	طعموا	لم يطعموا	المجموع
لم يصب	ا ١٩٢	ب ١١٣	٣٠٥
أصيب	حـ ٤	د ٣٤	٣٨
المجموع	٢٢٦	١١٧	٣٤٣

$$\therefore \text{معامل الاقتران} = \frac{١٩٢ \times ٣٤ - ١١٣ \times ٤}{١٩٢ \times ٣٤ + ١١٣ \times ٤}$$

$$= ٨٧ ر$$

وتدل هذه النتيجة على وجود علاقة طردية قوية بين التطعيم بالمصل وعدم الاصابة بالمرض . وقيمة هذا المعامل تكون دائماً أقل من ١ طالما كانت ب حـ أكثر من صفر ، أما إذا كانت تساوي صفر فان المعامل = ١ وفي هذه الحالة تكون هناك علاقة كاملة بين الظاهرتين التي نقوم ببحثها .

ونتيجة المثال السابق يمكن تحقيقها إذا حسبنا نسبة الذين طعموا ولم يصابوا بالمرض وهي ٩٨ ٪ تقريباً ، بينما نسبة الذين لم يطعموا ولم يصـــابوا تساوي ٧٦ ٪ تقريباً .

أما إذا كانت احدى الظاهرتين اللتين نبحث العلاقة بينهما أو كلتيهما تنقسم إلى أكثر من نوعين ، فان معامل الاقتران لا يساعدنا في هذه الحالة وعندئذ

نستخدم معامل التوافق الذي وضعه بيرسون لقياس العلاقة بين الصفات غير المقيسة ، أو بين صفات بعضها يقاس بالأرقام وبعضها لا يقاس .

ويحسب معامل التوافق من القانون التالي :

$$\text{معامل التوافق} = \sqrt{\frac{\text{ح} - ١}{\text{ح}}}$$

ولتطبيق هذا القانون نبوب البيانات الخاصة بالظاهرتين في جـــدول مزدوج ثم نربع كل تكرار في الجدول ونقسمه على حاصل ضرب التكرار الكلي العمودي في التكرار الكلي الافقي .

مثال ٥٨ :

البيانات التالية خاصة بعدد الطلبة حسب تقديرات نجاحهم في مادتين :

ص \ س	مقبول	جيد	جيد جدا	ممتاز	المجموع
مقبول	٢٧	٥٠	٣٠	٢١	١٢٨
جيد	٤٥	١٠٣	٨٢	٥٠	٢٨٠
جيد جدا	١٠	١٨	٢٣	١٤	٦٥
المجموع	٨٢	١٧١	١٣٥	٨٥	٤٧٣

السطر الأول : $\frac{١}{١٢٨}\left(\frac{٢٧^٢}{٨٢} + \frac{٥٠^٢}{١٧١} + \frac{٣٠^٢}{١٣٥} + \frac{٢١^٢}{٨٥}\right) = ٢٧٠٢ر$

السطر الثاني : $\frac{١}{٢٨٠}\left(\frac{٤٥^٢}{٨٢} + \frac{١٠٣^٢}{١٧١} + \frac{٨٢^٢}{١٣٥} + \frac{٥٠^٢}{٨٥}\right) = ٥٩٢٧ر$

السطر الثالث : $\frac{١}{٦٥}\left(\frac{١٠^٢}{٨٢} + \frac{١٨^٢}{١٧١} + \frac{٢٣^٢}{١٣٥} + \frac{١٤^٢}{٨٥}\right) = ١٤٣٧ر$

ح = ٢٧٠٢ر + ٥٩٢٧ر + ١٤٣٧ر = ١ر٠٠٦٦

$$\text{معامل التوافق} = \sqrt{\frac{١ر٠٠٦٦ - ١}{١ر٠٠٦٦}} = ٠٨ر$$

وتدلنا هذه النتيجة على عدم وجود علاقة بين هاتين الظاهرتين

الفصل الثاني عشر

التوضيح البياني للارتباط – الانحدار

إذا كان لدينا عدة أزواج من القيم غير المبوبة لظاهرتين فان خط الانحدار لظاهرة منهما على الأخرى هو الخط البياني الذي يمثل العلاقة بين الظاهرتين احداهما الظاهرة المستقلة ممثلة على المحور الافقي (Independent Variable) والاخرى الظاهرة الفرعية أو التابعة (Dependent Variable) ممثلة على المحور الرأسي . فان كان هناك علاقة بين الظاهرتين كان لهذا الخط وجود ، أما إذا انعدمت العلاقة لما أمكن رسم مثل هذا الخط ، بمعنى ان وجود العلاقة بينهما يؤدي إلى وجود اتجاه عام للنقط التي نحددها في الرسم أمام قيم الظاهرة المستقلة والقيم التابعة لها للظاهرة الاخرى . هذا الاتجاه العام إما أن يكون مستقيما فتكون العلاقة من الدرجة الاولى ويسمى الارتباط مستقيما ، واما أن يكون غير مستقيم فتكون العلاقة من درجة أعلى من الأولى ويسمى الارتباط في هذه الحالة غير مستقيم .

على انه يجب أن نلاحظ ما نقصده بالاتجاه العام حيث انما نعنيه بذلك لا يستلزم وجود جميع النقط على خط مستقيم أو غير مستقيم (خط الانحدار) وانما يكون هناك اتجاه عام لها يمكن تحديده بالرسم . على أنه إذا وقعت جميع النقط على خط الانحدار فان هذا دليل واضح على الارتباط الكامل بين الظاهرتين ، وكلما قربت النقط من خط الاتجاه العام أو وقع

معظمها عليه فان هذا دليل على شدة الارتباط بين الظاهرتين بينما إذا بعدت معظم النقط عن خط الانحدار فان هذا دليل على ضعف الارتباط بينهما . وبمعنى آخر كلما كان تشتت النقط حول خط الانحدار كبيراً كلما ضعف الارتباط والعكس كلما كان تشتتها ضعيفاً كان هذا دليل على شدة الارتباط بين الظاهرتين التي ندرس العلاقة بينهما .

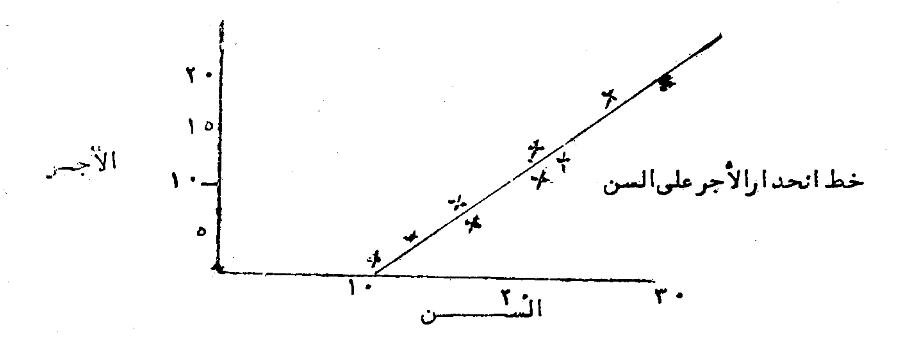
وخط الانحدار بذلك يوضح لنا هندسياً العلاقة بين الظاهرتين ، فهو يرينا كيف تميل الظاهرة التابعة إلى التغير نتيجة تغير معين في الظاهرة المستقلة . على أن هذا التغير في الظاهرة التابعة الذي يظهره الرسم ليس هو حتما نفس التغير الذي يحدث في الواقع العملي ، تماماً مثل المتوسط لعدة قيم ، فلا يعني هذا المتوسط أن جميع القيم متساوية حتما ولذا فخط الانحدار يسمى أحياناً خط العلاقة المتوسطة بين الظاهرتين حيث يعطينا القيمة النظرية للمتغير التابع التي تقابل قيمة معينة للمتغير المستقل . هذه القيمة الثالثة قد تكون هي نفسها القيمة الواقعية وقد تختلف عنها بعض الشيء . وكلما قربت القيم النظرية للمتغير التابع من قيمته الواقعية كلما كان ذلك دليلاً على أن خط الانحدار يمثل العلاقة بين الظاهرتين تمثيلاً صادقاً ، وكلما بعدت القيم النظرية عن القيم الواقعية كلما دل ذلك على ضعف تمثيل خط الانحدار للعلاقة بين الظاهرتين .

مثال ٥٩ :

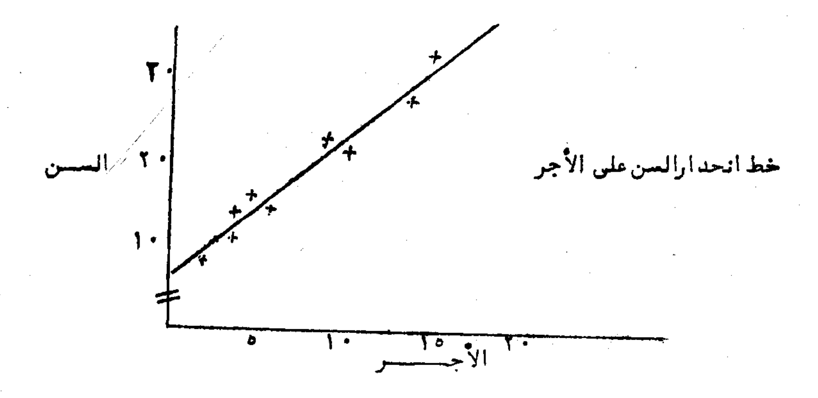
سن الطفل	الأجر اليومي بالقرش
١٢	٢
١٤	٣
٢١	٩
١٦	٦
٢٠	١١
٢٥	١٦
٣٠	١٩
١٥	٣
١٧	٤
٢٠	٧

لو مثلنا سن العامل على المحور الأفقي وأجره على المحور الرأسي ومثلنا كل زوج من القيم المتقابلة في الجدول السابق بنقطة بيانية لانتشرت النقط بنظام خاص يسمى بالاتجاه العام لها ، ولدلّ ذلك على نوع العلاقة بين الظاهرتين، ويمكن تمثيل هذه العلاقة بخط مستقيم كما فعلنا في الشكل. ويسمى الشكل الذي نحصل عليه باتباع هذه الخطوات في الرسم بشكل الانتشار (Scatter Diagram) والمستقيم الممثل له بخط الانحدار أو خط العلاقة المتوسطة .

وخط الانحدار السابق هو خط انحدار الأجر على السن حيث انه جرت العادة على وصف الخط باسم الظاهرة التابعة أولاً ويتضح للقارىء سبب هذه التسمية عند الكلام على فائدة خطوط الانحدار .



ويمكننا كذلك ان نعتبر ان الأجر هو المتغير المستقل فنرسمه على المحور الافقي وان السن هو المتغير التابع فنرسمه على المحور الرأسي وبذلك نحصل على خط انحدار السن على الأجر كما يتضح من الرسم التالي :



وفائدة خطوط الانحدار انها تمكن الباحث من تقدير قيمة المتغير التابع لو عرف قيمة المتغير المستقل ؛ فلو قيل لك ما الأجر الذي يجب أن يعطى لعامل عمره ٢٢ سنة مثلاً ، لأمكن استخدام الشكل الأول في استنباط القيمة

باقامة عمود على المحور الافقي عند السن ٢٢ سنة فيقابل خط الانحدار في نقطة يكون احداثيها الرأسي هو القيمة المطلوبة (١١ر٠٩ قرشاً) . وبالمثل يمكن استخدام خط انحدار السن على الأجر لتحديد ما نتوقع أن يكون عليه عمر العامل ذي أجر معين .

وتحديد القيم التابعة على أساس خطوط الانحدار بطريقة بيانية تعوزها الدقة وتعتمد الى حد كبير على الحكم الشخصي للباحث نفسه حيث عليه يتوقف رسم خط الانحدار وهو أمر ليس من السهل ولا يمكن ان يتفق فيه اثنان الا نادرا . لهذا يجب تمثيل العلاقة بطريقة جبرية دقيقة .

طرق ايجاد معادلة الانحدار :

لايجاد معادلة الانحدار يجب أولاً رسم شكل الانتشار حتى نتبين ما اذا كانت العلاقة بين الظاهرتين من الدرجة الاولى اي يصورها خط مستقيم أو من درجة اعلى أي يصورها خط غير مستقيم .

فاذا كانت العلاقة من الدرجة الاولى فنعمل على ايجاد معادلة المستقيم من واقع القيم المتناظرة التي تتخذها الظاهرتين ، وذلك إما بطريقة النضمين إذا تبين لنا ان جميع النقط التي تمثل القيم المتناظرة تقع على خط مستقيم ، أما إذا كانت النقط تنتشر حول خط مستقيم فيمكن ايجاد المعادلة اما بطريقة أصغر المربعات وإما بطريقة تعتمد على كون العلاقة بين الظاهرتين تتوقف على درجة تشتت النقط حول خط العلاقة المتوسطة (خط الانحدار) .

مثال ٦٠ :

ايجاد معادلة خط انحدار $\frac{ص}{س}$ من واقع القيم الآتية :

(١) التضمين —

س	ص	
١	٥	
٢	٧	٢
٣	٩	٢
٤	١١	٢
٥	١٣	٢

(استخدمنا طريقة التضمين حيث اذا رسمنا شكل الانتشار تقع جميع النقط على خط مستقيم)

ص = ٥ + ٢ (س − ١)

٥ + ٢س − ٢

= ٢س + ٣ وهي المعادلة المطلوبة .

(٢) بتطبيق طريقة أصغر المربعات

ص = م س + ب

مجـ ص = م مجـ س + ن ب (١)

مجـ س ص = م مجـ س^{٢} + ب مجـ س (٢)

وبالتعويض في هاتين المعادلتين يمكن أن نحسب قيمة كل من م ، ب

٤٠ = ١٥م + ٥ ب (٣)

١٥٥ = ٥٥م + ١٥ب (٤)

وبضرب المعادلة −١− في ٣ تكون

١٣٥ = ٤٥م + ١٥ ب

١٥٥ = ٥٥م + ١٥ ب

وبالطرح ٢٠ = ١٠م

م = ٢

س	ص	س^{٢}	س ص
١	٥	١	٥
٢	٧	٤	١٤
٣	٩	٩	٢٧
٤	١١	١٦	٤٤
٥	١٣	٢٥	٦٥
١٥	٤٥	٥٥	١٥٥

وبالتعويض في المعادلة رقم – ٣ – يكون :

٤٥ = ٣٠ + ٥ ب

١٥ = ٥ ب

ب = ٣

ص = ٢ س + ٣ وهي المعادلة المطلوبة .

وبالمثل يمكن ايجاد معادلة س على ص حيث نستخدم الصيغة س = م⁻ ص + ب⁻ .

(٣) على أساس معرفة الوسط الحسابي لكل من الظاهرتين وانحرافها المعياري ودرجة الارتباط بينهما (معامل الارتباط) يمكن ايجاد معادلة الانحدار كالآتي :

معادلة انحدار $\frac{ص}{س}$ هي (ص – ص⁻) = ر $\frac{ع ص}{ع س}$ (س – س⁻)

ومعادلة انحدار $\frac{س}{ص}$ هي (س – س⁻) = ر $\frac{ع س}{ع ص}$ (ص – ص⁻)

ومن هاتين المعادلتين نعرف ان معامل انحدار $\frac{ص}{س}$ = ر $\frac{ع ص}{ع س}$

وهو يساوي ظل زاوية انحدار خط انحدار $\frac{ص}{س}$ ومعامل انحدار $\frac{س}{ص}$ = ر $\frac{ع س}{ع ص}$ وهو يساوي ظل زاوية انحدار خط انحدار $\frac{س}{ص}$. فاذا علم لدينا معامل انحدار $\frac{ص}{س}$ ومعامل انحدار $\frac{س}{ص}$ نستطيع بضربهما وايجاد الجذر

للناتج أن نحصل على معامل الارتباط حيث ر $\frac{ع_ص}{ع_س}$ × ر $\frac{ع_س}{ع_ص}$ = $ر^٢$

فاذا رجعنا للقيم السابقة امكننا ايجاد معادلة انحدار $\frac{ص}{س}$ كالآتي :

س	ص	$س^٢$	$ص^٢$
١	٥	١	٢٥
٢	٧	٤	٤٩
٣	٩	٩	٨١
٤	١١	١٦	١٢١
٥	١٣	٢٥	١٦٩
١٥	٤٥	٥٥	٤٤٥

$\bar{س} = \frac{١٥}{٥} = ٣$ ، $\bar{ص} = \frac{٤٥}{٥} = ٩$

$ع_س = \sqrt{\frac{٥٥}{٥} - ٩} = \sqrt{٢}$

$ع_ص = \sqrt{\frac{٤٤٥}{٥} - ٨١} = \sqrt{٨٩ - ٨١} = \sqrt{٨}$

ان جميع النقط تقع على خط مستقيم يكون الارتباط كامل كما أشرنا إلى ذلك من قبل .

وبذلك تكون معادلة انحدار $\frac{\text{ص}}{\text{س}}$ هي

$$\text{ص} - ٩ = ١ \times \frac{\sqrt{٨}}{\sqrt{٢}} \; (\text{س} - ٣)$$

$$\therefore \; \text{ص} - ٩ = ٢\,\text{س} - ٦$$

$$\therefore \; \text{ص} = ٢\,\text{س} + ٣$$ وهي المعادلة المطلوبة

وبالمثل يمكن ايجاد معادلة انحدار $\frac{\text{س}}{\text{ص}}$

طرق ايجاد معادلة الانحدار (من الدرجة الثانية)

إذا تبين من شكل الانتشار ان المعادلة من الدرجة الثانية أي ان خط الانحدار يكون منحني نستطيع ايجاد معادلة الانحدار اما بطريقة التضمين إذا كانت جميع النقط تقع على المنحنى أو بطريقة أصغر المربعات إذا كانت تنتشر حوله .

مثال ٦١

(١) بالتضمين —

س	ص		
١	٣		
		٤	
٢	٧		١
		٦	
٣	١٣		١
		٨	
٤	٢١		١
		١٠	
٥	٣١		

$$\text{ص} = ٣ + ٤\,(\text{س} - ١) + ١\,(\text{س} - ١)(\text{س} - ٢)$$

$$= ٣ + ٤\,\text{س} - ٤ + \text{س}^{٢} - ٣\,\text{س} + ٢$$

$$= \text{س}^{٢} + \text{س} + ١$$ وهي المعادلة المطلوبة

(٢) بطريقة أصغر المربعات

ص = ا س^{٢} + ب س + حـ   المعادلة العامة للمنحنى

: مجـ ص = ا مجـ س^{٢} + ب مجـ س + ن حـ   — ١ —

، مجـ س ص = ا مجـ س^{٣} + ب مجـ س^{٢} + حـ مجـ س   — ٢ —

، مجـ س^{٢} ص = ا مجـ س^{٤} + ب مجـ س^{٣} + حـ مجـ س^{٢} — ٣ —

من هذه المعادلات الثلاث نستطيع ان نحسب قيمة ا ، ب ، حـ

س	ص	س ص	س^{٢}	س^{٣}	س^{٤}	س^{٢}ص
١	٣	٣	١	١	١	٣
٢	٧	١٤	٤	٨	١٦	٣٨
٣	١٣	٣٩	٩	٢٧	٨١	١١٧
٤	٢١	٨٤	١٦	٦٤	٢٠٦	٢٣٦
٥	٣١	١٥٥	٢٥	١٢٠	٦٢٥	٧٧٥
١٠	٧٥	٢٩٥	٥٥	٢٢٥	٩٧٩	١٢٥٩

: ٧٥ = ٥٥ ا + ١٥ ب + ٥ حـ   — ١ —

، ٢٧٥ = ٢٢٥ ا + ٥٥ ب + ١٥ حـ   — ٢ —

، ١٢٥٩ = ٩٧٩ ا + ٢٢٥ ب + ٥٥ حـ   — ٣ —

بضرب المعادلة — ١ — في ٣ ينتج :

٢٢٥ = ١٦٥ ا + ٤٥ ب + ١٥ حـ

٢٩٥ = ٢٢٥ ا + ٥٥ ب + ١٥ حـ

بالطرح ينتج ٧٠ = ٦٠ ا + ١٠ ب   — ٤ —

وبضرب المعادلة - ١ - في ٣ ينتج

٨٢٥ = ٦٠٥ ا + ١٦٠ ب + ٥٥ حـ

١٢٥٩ = ٩٧٩ ا + ٢٢٥ ب + ٥٥ حـ

بالطرح ينتج ٤٣٤ = ٤٣٤ ا + ٦٠ ب — ٥ —

بضرب المعادلة ٤ في ٦ ينتج :

٤٢٠ = ٣٦٠ أ + ٦٠ ب

٤٣٤ = ٣٧٤ أ + ٦٠ ب

بالطرح ينتج ١٤ = ١٤ أ

: أ = ١

وبالتعويض في المعادلة ٤ عن قيمة أ ينتج :

٧٠ = ٦٠ + ١٠ ب

١٠ = ١٠ ب

: ب = ١

وبالتعويض في المعادلة - ١ - عن قيمة أ ، ب ينتج :

٧٥ = ٥٥ + ١٥ + ٥ حـ

: ٧٥ = ٧٠ + ٥ حـ

: ٥ = ٥ حـ

: حـ = ١

المعادلة هي ص = س^{٢} + س + ١ وهي المعادلة المطلوبة .

ايجاد معادلة الانحدار في حالة القيم المبوبة

لا تختلف خطوات العمل عند ايجاد معادلة الانحدار في حالة القيم المبوبة عن الخطوات السابقة الا في أنه يجب أولاً تحويل البيانات الواردة في الجدول المزدوج الى قيم بسيطة لكل من الظاهرتين . وذلك بأن نأخذ مراكز فئات المتغير المستقل ونعتبرها قيماً لهذا المتغير ثم نبحث عن القيم التابعة بحساب المتوسط المقابل لكل قيمة من قيم المتغير المستقل من واقع الجدول المزدوج .

بعد ذلك يصبح لدينا قيم بسيطة لكل من المتغيرين فنرسم شكل الانتشار ومنه نحدد درجة المعادلة ، فاذا كان الاتجاه العام مستقيماً عملنا على ايجاد معادلة المستقيم ص = م س + ب بالطرق التي سبق مناقشتها واذا كان منحنى عملنا على ايجاد معادله الدرجة الثانية ص = أ س^{٢} + ب س + ح

مثـال ٦٢ :

الآتي بيان بتوزيع تكراري مزدوج لايجار ١٠٠ موظف والمرتبات التي يتقاضونها . المطلوب ايجاد معادلة انحدار الايجار على المرتب .

المرتب / الايجار	٣-	٥-	٧-	٩-	١١-	المجموع
صفر-	٩	٢٩				١٨
١-	١	١١	١٢			٢٤
٢-			١٣	٥		١٨
٣-				١٥	٥	٢٠
المجموع	١٠	٤٠	٢٥	٢٠	٥	١٠٠

الحـــل :

خط انحدار ايجار المسكن على المرتب هو الخط البياني الذي يمثل العلاقة

بين مراكز فئات المرتب (المتغير المستقل) ومتوسط الايجار الذي يدفعه أفراد كل فئة (المتغير التابع) ويمكن حساب هذا المتوسط بضرب مركز كل فئة لايجار المسكن في التكرار المقابل له في الخانة المواجهة لمركز فئة المتغير المستقل .

فمتوسط الايجار المقابل للمرتب ٤ (مركز الفئة ٣ — ٥) =

(٥ر × ٩ + ١ر٥ × ١) / ١٠ = ٦ر

ومتوسط الايجار المقابل للمرتب ٦ (مقابل الفئة ٥ — ٧) =

(٥ر × ٢٩ + ١ر٥ × ١١) / ٤٠ = ٧٨ر

ومتوسط الايجار المقابل للمريب ٨ (مركز الفئة ٧ — ٩) =

(١ر٥ × ١٣ + ٢ر٥ × ١٣) / ٢٥ = ٢٠ر

ومتوسط الايجار المقابل للمرتب ١٠ (مركز الفئة ٩ - ١١) =

(٢ر٥ × ٥ + ٣ر٩ × ١٥) / ٢٠ = ٢٥ر

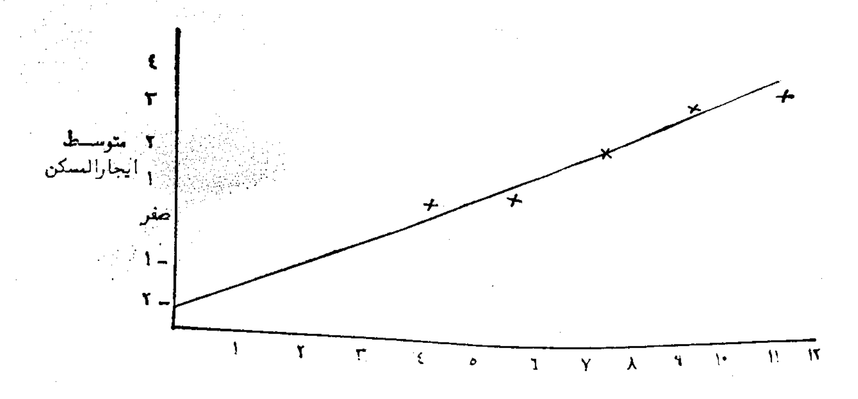
ومتوسط الايجار المقابل للمرتب ١٢ (مركز الفئة ١١ — ١٣) =

(٣ر٥ × ٥) / ٥ = ٣ر٥

وبذلك يمكن وضع قيم المتغيرين في الشكل الآتي :

المرتب	متوسط ايجار المسكن
٤	٦ر
٦	٧٨ر
٨	٢ر٠٢
١٠	٣ر٢٥
١٢	٣ر٥

بعد ذلك نرسم هذه البيانات في شكل بياني كالآتي :



المرتب

من شكل الانتشار يتبين ان الاتجاه العام للعلاقة بين هذين المتغيرين هو مستقيم فالمعادلة التي تصور العلاقة بينهما تكون من الدرجة الاولى . وحيث ان جميع النقط لا تقع على خط الانحدار لهذا لا يمكن ايجاد المعادلة بطريقة التضمين . وبذلك لا يمكن ايجاد المعادلة الا بطريقة اصفر المربعات أو باستخدام الصيغة $ص - \bar{ص} = ر \frac{ع ص}{ع س} (س - \bar{س})$. وخط انحدار المرتب على ايجار المسكن هو الخط البياني الذي يمثل العلاقة بين مراكز فئات الايجار (المتغير المستقل) ومتوسطات المرتب المقابلة لهذه المراكز .

ومن الجدول السابق يمكن ايجاد متوسطات المرتب المقابلة لمراكز فئات الايجار كالآتي :

متوسط المرتب المقابل للايجار ٥ر (مركز الفئـــة صفر – ١) =

$$\frac{٢٩ \times ٦ + ٩ \times ٤}{٣٨} = ٥ر٥٣$$

متوسط المرتب المقابل للايجار ٥ر١ (مركز الفئة ١ - ٢) =

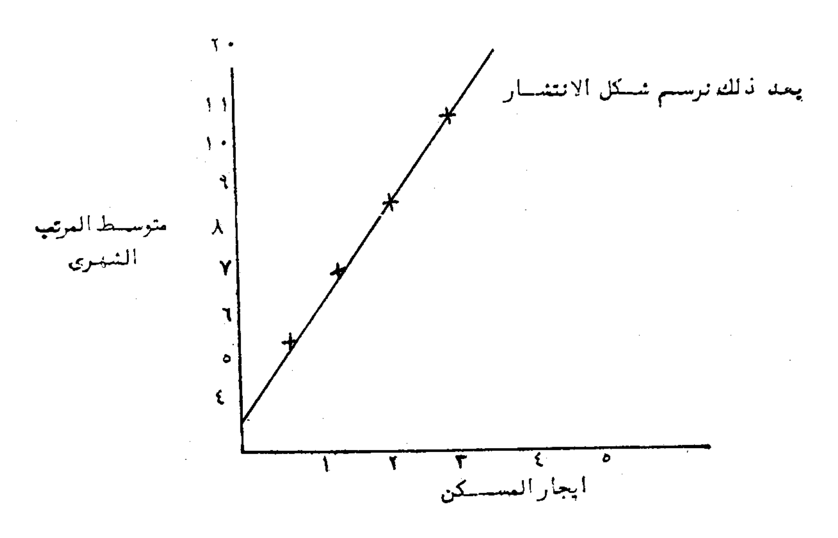
$$\frac{١٢\times ٨ + ١١\times ٦ + ١\times ٤}{٢٤} = ٩٢ر$$

» » » » ٥ر٢ (مركز الفئة ٢ – ٣) =

$$\frac{٥ = ١٠ + ١٣ \times ٨}{١٨} = ٨ر٥٥ .$$

» » » » ٥ر٣ (مركز الفئة ٣ – ٤) =

$$\frac{٥ \times ١٢ + ١٥\times ١٠}{٢٠} = ١٠ر٥ .$$



من شكل الانتشار يتبين أن المعادلة من الدرجة الأولى حيث يظهر أن الاتجاه العام للعلاقة بين المتغيرين مستقيم . ولا يمكن استعمال طريقة التضمين لايجاد لمعادلة لأن جميع النقط ليست على خط الاتجاه العام ، ولذلك نستعمل اما طريقة أصغر المربعات أو الطريقة التي تعتمد على الصيغة :

$$ص - \bar{ص} = ر \frac{ع_ص}{ع_س} (س - \bar{س}) .$$

ايجاد معامل الارتباط بمعلومية الانحدار :

يتبين لنا مما سبق أن معامل الارتباط =

$$\sqrt{\frac{\text{معامل انحدار ص}}{س} \times \frac{\text{معامل انحدار س}}{ص}}$$

ولهذا اذاعلم لدينا كل من هذين المعاملين عن طريق معرفة معادلتي الانحدار ص = م س + ب ، س = $\bar{م}$ ص + $\bar{ب}$ يمكن حساب معامل الارتباط بايجاد الجذر التربيعي لمعاملي الانحدار وهما م ، $\bar{م}$ في هاتين المعادلتين .

وبذلك ر = $\sqrt{م \times \bar{م}}$

وقد يوجد لدينا خطي الانحدار فيمكن حساب معاملي الانحدار عن طريق حساب ميل كل خط من الخطين . وميل خط الانحدار هو ظل زاوية الانحدار ، ويساوي $\frac{\text{المقابل}}{\text{المجاور}}$.

مثـــال ٦٣ :

من الرسم التالي الذي يظهر فيه خطي الانحدار ، احسب معامل الارتباط :

ميل خط الانحدار ص = م س + ب

= ظل زاوية الانحدار م ع س = $\frac{\text{م س}}{\text{س ع}}$

ميل خط الانحدار س = م ص + ب

= ظل زاوية الانحدار س م ص = $\frac{\text{س ص}}{\text{م س}}$

∴ ر = $\sqrt{\frac{\text{م س}}{\text{س ع}} \times \frac{\text{س ص}}{\text{م س}}} = \sqrt{\frac{\text{س ص}}{\text{س ع}}}$

الخطأ المعياري :

ذكرت فيما سبق ان العلاقة بين المتغيرين يمثلها خط الانحدار الذي يمكن تسميته بخط العلاقة المتوسطة . كذلك ذكرت ان النقط في شكل الانتشار

لا تقع جميعها في أغلب الأحيان على خط الانحدار وانما تنتشر حوله ، وان خط العلاقة المتوسطة يمكن تشبيهه بالوسط الحسابي لعدّة قيم . فالوسط الحسابي يمثل مجموعة القيم ويكون مجموع مربع انحرافات هذه القيم عنه أقـــل من مجموع مربـع انحرافات القيم عن أي قيمة أخرى غيره (كذلك خط العلاقة المتوسطة فهو يمثل النقط المنتشرة حوله على أساس أن مجموع مربع انحرافات النقط عنه أقـــل من مجموع مربـع انحرافاتها عن أي خط آخر) وهذا هو السبب في تسمية طريقة ايجاد معادلته بطريقة أصغر المربعات. وقد رأينا فيما سبق كيف استخدمت هذه الخاصية للوسط الحسابي في قياس مدى تشتت القيم حوله (الانحراف المعياري) . كذلك يمكن استخدام هذه الخاصية لخط العلاقة المتوسطة في قياس مدى تشتت النقط حوله. والخطأ المعياري هو الذي يقيس هذا التشتت ويرمز له بالرمز ر سواء كان الخطأ المعياري الذي نحسبه هو لمعـــادلة انحدار $\frac{ص}{س}$ أو $\frac{س}{ص}$.

ويمكن تعريف الخطأ المعياري بأنه المقياس الذي يقيس لنا درجة دقة معادلة الانحدار في تمثيل العلاقة بين المتغيرين تمثيلاً جبرياً ، وبالطبـع كلما صغر الخطأ المعياري كلما دل ذلك على أن المعادلة تمثل العلاقة تمثيلا صادقاً والعكس كلما كبر الخطأ المعياري نقصت درجة دقة المعادلة في تصويرها للعلاقة . وبمعنى آخر تكون القيم التابعة المستنبطة من معادلة الانحدار قيما يمكن الاعتماد عليها إذا صغر الخطأ المعياري للمعادلة ولا يمكن الاعتماد عليها إذا كبر الخطأ المعياري .

والخطأ المعياري هو الجذر التربيعي لمتوسط مجموع مربـع انحرافات القيم النظرية للمتغـــير ص أو المتغير س عن قيمه الواقعية . والقيم النظرية هي المستنبطة من معاملة الانحدار أو من خط العلاقة المتوسطة بالرسم البياني .

ويمكن حساب الخطأ المعياري لمعادلة الدرجة الأولى كالآتي :

خ^{2} لمعادلة انحدار $\frac{\text{ص}}{\text{س}} = \frac{1}{\text{ن}}$ مج ص^{2} − (م مج س ص + ب مج ص)

خ^{2} لمعادلة انحدار $\frac{\text{س}}{\text{ص}} = \frac{1}{\text{ن}}$ مج س^{2} − (م− مج س ص + ب− مج س)

أما إذا كانت معادلة الانحدار من الدرجة الثانية فيمكن حساب الخطأ المعياري كالآتي :

ح^{2} لمعادلـــة انحدار $\frac{\text{ص}}{\text{س}}$ =

$\frac{1}{\text{ن}}$ مج ص^{2} − (أ مج س^{2} ص + ب مج س ص + حـ مج ص)

ح^{2} لمعادلـــة انحدار $\frac{\text{س}}{\text{ص}}$ =

$\frac{1}{\text{ن}}$ مج س^{2} - (أَ مج ص^{2} س + ب− مج س ص + حَ مج ص)

في هذه المعاملات نلاحظ أن

ن = عدد النقط .

مج ص^{2} = مجموع مربع القيم الواقعية للمتغير ص .

مج س^{2} = مجموع مربع القيم الواقعية للمتغير س

القيم الواقعة داخل الأقواس الصغيرة هي مجموع مربع القيم النظرية لكل من ص ، س كل حسب معادلة الانحدار الخاص به .

الخطأ المعياري والارتباط :

أشرت سابقاً إلى أنه اذا وقعت جميع النقط على خط الانحدار فان

$ر^2 = 1 - \frac{ح^2_ص}{ع^2_ص}$ ، أما إذا علم لدينا الخطأ المعياري لمعادلة انحـــدار $\frac{س}{ص}$

فتكون $ر^2 = 1 - \frac{ح^2_س}{ع^2_س}$. هذه العلاقة الجبرية بين معامل الارتباط س والخطأ المعياري $ح^2$ والانحراف المعياري للقيم التابعة يمكن إثباتها .

والعلاقة السابقة بين معامل الارتباط والخطأ المعياري صحيحة كذلك إذا كانت الانحدار من درجة أعلى من الأولى ، والفرق الوحيد هو لمجرد التمييز بين الارتباط المستقيم وغير المستقيم فنسمي مقياس الارتباط غير المستقيم دليل الارتباط ويرمز له بالرمز ط وهو يستعمل في قياس الارتباط غير المستقيم من واقع قيم غير مبوبة .

وبذلك يكون $ط^2 = 1 - \frac{خ^2_ص}{ع^2_ص}$ في حالة انحدار $\frac{ص}{س}$

، $ط^2 = 1 - \frac{خ^2_س}{ع^2_س}$ في حالة انحدار $\frac{س}{ص}$

ونلاحظ ان القيمة التي نحسبها لمعامل الارتباط (ر) ولدليل الارتباط (ط) لا تتساوى في حالة الارتباط غير المستقيم – بمعنى انه إذا حسبنا ر لقياس الارتباط بين متغيرين علاقتهما من الدرجة الثانية (باستخدام قانون بيرسون مثلا) فانها تكون مقيـــاس غير دقيق للارتباط بين هذين المتغيرين حيث ان ر تحسب على أساس ان العـــلاقة من الدرجة الأولى والمقياس الاصح هو ط الذي يحسب من واقع المعادلات المذكورة أعلاه .

لذلك يرى بعض الاحصائيين عدم استعمال ر لقيـــاس الارتباط بين متغيرين ويقترحون استخدامه فقط في قياس درجة استقامة خطوط الانحدار

وتسميته تبعاً لذلك بمعامل الاستقامة وذلك لأنه إذا كان ر كبيراً كان مقدار $خ^2$ صغيراً مما يدل على ان الخط المستقيم يوافق جيدا النقط المعطاة ، أي ان خط الانحدار قريب من المستقيم . أما إذا كان ر صغيراً كان مقدار $خ^2$ كبيراً حول الخط المستقيم مما يدل على ان هذا الخط المستقيم لا يوافق هذه النقط جيداً ، أي ان خط الانحدار بعيد عن الاستقامة .

كما اننا إذا حسبنا ر ثم حسبنا ط عند قياس العلامة بين متغيرين وكانت ر أكبر من ط كان ذلك دليلاً قوياً على ان القيم المعطـاة يوافقها مستقيم أفضل من أي خط آخر . والعكس إذا كانت ط أكبر من ر دل ذلك على أن القيم المعطاة يوافقها منحنى أفضل من خط المستقيم ولذلك كان تشتتها حول الخط من الدرجة الثانية مثلاً الذي استخدمناه في حساب ط أقل من تشتتها حول الخط المستقيم الذي افترضناه موافقاً لتمثيـل هذه القيم عند حساب المعامل ر لها .

ومن ناحية اخرى إذا رسمنا شكل الانتشار وتبين لنـــا ان خط الاتجاه العام مستقيم يمكن حساب ر ، أما إذا كان منحني فيجب حساب ط بايجاد المعادلة ثم حساب الخطأ المعياري لها .

على اننا نعرف ان معرفة الاتجاه العام للعلاقة بين متغيرين ليس من السهل تحديدها بدقّة فهي تعتمد إلى حد كبير على الرأي الشخصي وبهذه الطريقة يدخل العامل الشخصي في أبحاثنا ويتحكم في المقياس الذي نحصل عليه لدرجة العلاقة بين المتغيرين تحت البحث .

نسبة الارتباط :

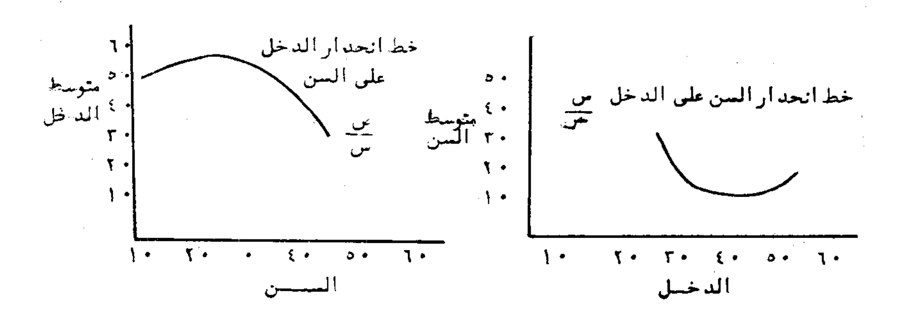
عندما تكون قيم المتغيرين مبوبه في جدول تكراري مزدوج وتبين لنا ان العلاقة بينهما ليست من الدرجة الأولى فلا يمكن حساب ر كمقياس دقيق للعلاقة بين المتغيرين . كذلك يكون من الصعب حساب ط لأن هذا لا بد

له من توفيق منحنى يصور جبريا العلاقة بين المتغيرين (أصغر المربعات) . لذلك وضع بيرسون مقياساً سماه نسبة الارتباط لاستخدامه في مثـــل هذه الأحوال ويرمز له بالحرف ي .

ونسبة الارتباط بين أي متغيرين = $\frac{ع_ص}{ع_ص}$ حيـــث $ع_ص$ تـــدل على الانحراف المعياري لتوسطات قيم ص المقابلة لمراكز فئات س ، ع ص هي الانحراف المعياري لقيم ص نفسها . كذلك فان ي = $\frac{ع_س}{ع_س}$ حيث $ع_س$ الانحراف المعياري لمتوسطات قيم س المقابلة لمراكز فئات ص ، $ع_س$ هي الانحراف المعياري لقيم س .

مثال ٦٤ :

السن / الدخل	صفر –	١٠ –	٢٠ –	٣٠ –	٤٠ –	المجموع	متوسطات العمر
٣٠ –	٤				٨	١٢	٢١ر٦٧
٣٥ –	٨				٤	١٢	١٨ر٣٢
٤٠ –	١٤	١٢	١٠	٨		٤٤	١٧ر٧٣
٤٥ –		٦	١٤			٢٠	٢٢ر٠٠
٥٠ –			١٢			١٢	٢٥ر٠٠
المجموع	٢٦	١٨	٣٦	٨	١٢	١٠٠	
متوسطات الدخل	٣٩ر٤٢	٤٤ر١٧	٤٧ر٧٨	٤٢ر٥٠	٣٤ر١٧		



يتضح من هذا الرسم ان العلاقة بين هذين المتغيرين ليست من الدرجة الاولى وحيث أن القيم الخاصة بهما مبوبة فتكون نسبـــة الارتباط ى ٥ ي أحسن مقياس لدرجة العلاقة بينهما . ولحساب ى نتبع الخطوات الآتية :

اولا — حســاب ع

متوسطات الدخل	ك	متوسطات الدخل في ك	ح من الوسط الحسابي	ح ك	ح² ك
٣٩ر٤٢	٢٦	١٠٢٤ر١٢	—٣ر٤٨	—٩٠ر٤٨	٣١٤ر٨٧
٤٤ر١٧	١٨	٧١٥ر٠٦	+١ر٢٧	+٢٢ر٨٦	٢٩ر٠٢
٤٧ر٧٨	٣٦	١٧٢٠ر٠٨	+٤ر٨٨	+١٧٥ر٦٨	٥٨٧ر٣٢
٤٢ر٥٠	٨	٣٤٠ر٠٠	—ر٤٠	—٣ر٢٠	١ر٢٨
٣٤ر١٧	١٢	٤١٠ر٠٤	—٨ر٧٣	—١٠٤ر٧٦	٩١٤ر٥٥٠
المجموع	١٠٠	٤٢١٠ر١			١٨٤٧ر٠٥

$\bar{س}$ (الوسط الحسابي) $= \frac{٤٣١٠ر١}{١٠٠} = ٤٢ر١٠١$

$$ع = \sqrt{\frac{١٨٤٧ر٠٥}{١٠٠}} = \sqrt{١٨ر٤٧٠٥} = ٤ر٢٩٨٠$$

ثانيا — حساب ع س

الفئات	ك	م	ك م	ح عن الوسط الحسابي	ح ك	ح² ك
—٣٠	١٢	٣٢ر٥	٣٩٠	—١٠ر٤	—١٢٤ر٨	١٢٩٧ر٩١
—٣٥	١٢	٣٧ر٥	٤٥٠	—٥ر٤	—٦٤ر٨	٣٤٩ر٩٢
—٤٠	٤٤	٤٢ر٥	١٨٧٠	—ر٤	—١٧ر٦	٧ر٠٤
—٤٥	٢٠	٤٧ر٥	٩٥٠	+٤ر٦	+٩٢ر٠	٤٢٣ر٢٠
—٥٠	١٢	٥٢ر٥	٦٣٠	+٩ر٦	+١١٥ر٢	١١٠٥ر٦٢
المجموع	١٠٠		٤٢٩٠			٣١٨٤

$$\bar{س} = \frac{٤٢٩٠}{١٠٠} = ٤٢ر٩ \text{ .}$$

$$ع_{س} = \sqrt{\frac{٣١٨٤}{١٠٠}} = \sqrt{٣١٫٨٤} = ٥٫٦٤٣$$

$$\text{نسبة الارتباط ي} = \frac{٤٫٢٩٨}{٥٫٦٤٣} = ٠٫٧٦٢$$

ومما يجب التنويه به ان نسبة الارتباط مقياس أعم لحساب الارتباط بين ظاهرتين سواء كانت العلاقة بينهما مستقيمة أو غير مستقيمة ، بخلاف معامل الارتباط الذي لا يجب أن نستعمله الا إذا تأكدنا أن الانحدار بين المتغيرين مستقيم أو قريب جداً من الاستقامة .

استنباط قوانين خطوط الانحدار من الدرجة الاولى :

عندما يكون الانحدار مستقيماً تكون المعادلة التي تصور العلاقة بين الظاهرتين تصويراً جبرياً هي :

$$ص = م\ س + ب \qquad (١)$$

$$\text{أو} \quad س = م^{-}\ ص + ب^{-} \qquad (٢)$$

ولايجاد المعادلة الأولى نتبع طريقة المربعات الصغرى حيث نستخدم المعادلتين الآتيتين لايجاد قيمة كل من م ، ب :

$$مج\ ص = م\ مج\ س + ن\ ب \qquad (٣)$$

$$مج\ س\ ص = م\ مج\ س^{٢} + ب\ مج\ س \qquad (٤)$$

ونلاحظ أنه يمكن وضع المعادلة (٣) في الشكل التالي :

$$ن\ \bar{ص} = ن\ م\ \bar{س} + ن\ ب$$

بالقسمة على ن ينتج ان $\bar{ص} = م\ \bar{س} + ب$

$$\therefore ب = \bar{ص} - م\ \bar{س} \qquad (٥)$$

كذلك يمكن التعويض في المعادلة (٤) عن قيمة مجـ س ص من القانون الثاني للارتباط (بيرسون) ، وعن $\text{س}_{\text{م}}$ من قانون الانحراف المعياري .

$$\text{ر} = \frac{\text{مجـ س ص} - \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}}}{\text{ن}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}}}$$

$$\therefore \text{مجـ س ص} = \text{ن ر}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}} + \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}}$$

$$،\ \text{ع}^{٢}_{\text{س}} = \frac{\text{مجـ س}^{٢}}{\text{ن}} - \bar{\text{س}}^{٢}$$

$$\therefore \text{ع}^{٢}_{\text{س}} + \bar{\text{س}}^{٢} = \frac{\text{مجـ س}^{٢}}{\text{ن}}$$

$$\therefore \text{مجـ س}^{٢} = \text{ن}\ \text{ع}^{٢}_{\text{س}} + \text{ن}\ \bar{\text{س}}^{٢}$$

.. بالتعويض في المعادلة (٤) نحصل على : –

$$\text{ن ر}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}} + \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}} = \text{ن م}\ \text{ع}^{٢}_{\text{س}} + \text{ن م}\ \bar{\text{س}}^{٢} + \text{ن}\ \bar{\text{س}}\ (\bar{\text{ص}} - \text{م}\ \bar{\text{س}})$$

$$\therefore \text{ن ر}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}} + \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}} = \text{ن م}\ \text{ع}^{٢}_{\text{س}} + \text{ن م}\ \bar{\text{س}}^{٢} + \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}} - \text{ن م}\ \bar{\text{س}}^{٢}$$

$$\therefore \text{ن ر}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}} = \text{ن م}\ \text{ع}^{٢}_{\text{س}}$$

$$\therefore \text{م} = \frac{\text{ن ر}\ \text{ع}_{\text{س}}\ \text{ع}_{\text{ص}}}{\text{ن}\ \text{ع}^{٢}_{\text{س}}}$$

$$\therefore م = ر \frac{ع_{ص}}{ع_{س}} \qquad (٦)$$

بالتعويض بقيمة ب وقيمة م في معادلة الانحدار $ص = م\, س + ب$

$$\therefore ص = ر \frac{ع_{ص}}{ع_{س}} س + \bar{ص} - ر \frac{ع_{ص}}{ع_{س}} \bar{س}$$

$$\therefore ص - \bar{ص} = ر \frac{ع_{ص}}{ع_{س}} (س - \bar{س}) \qquad (٧)$$ وهي المعادلة

التي نستعملها لإيجاد انحدار $\frac{ص}{س}$ عندما يكون معلوماً لدينا المقاييس $\bar{س}$ ، $\bar{ص}$ ، $ع_{ص}$ ، $ع_{س}$ ، ر خاصة عندما تكون القيم مبوبة في توزيع تكراري مزدوج

وبالمثل يمكن استنباط معادلة انحدار $\frac{س}{ص}$ وهي $س - \bar{س} = ر \frac{ع_{س}}{ع_{ص}} (ص - \bar{ص})$

قانون الخطأ المعياري لمعادلة الانحدار من الدرجة الاولى :

نفترض أن لدينا ظاهرتان س ، ص وان قيمها المتناظرة هي كالآتي :

$س_1$	$ص_1$
$س_2$	$ص_2$
$س_3$	$ص_3$
$س_4$	$ص_4$
..	...
$س_ن$	$ص_ن$

واننا وجدنا معادلة انحدار $\frac{ص}{س}$ فكانت ص = م س + ب وبذلك إذا عوضنا عن س بقيمها نحصل على القيم النظرية للمتغير ص التي تقابل قيمه الواقعية ، وتكون هذه القيم النظرية هي م $س_1$ + ب ، م $س_2$ + ب ، م $س_3$ + ب ، م $س_4$ + ب ، .. . م $س_ن$ + ب .

وبذلك يكون الفرق بين كل قيمة واقعية للمتغير ص وقيمته النظرية المقابلة هي :

$ص_1$ − (م $س_1$ + ب) ، $ص_2$ − (م $س_2$ + ب) ، $ص_3$ − (م $س_3$ + ب) ، $ص_4$ − (م $س_4$ + ب)، . $ص_ن$ − ($س_ن$ +ب)

وبتربيع هذه الفروق نحصل على : —

١) $ص_1^2$ − ٢م $س_1$ $ص_1$ + $م^2$ $س_1^2$ − ٢ب $ص_1$ + ٢م ب $س_1$+$ب^2$

٢) $ص_2^2$ − ٢م $س_2$ $ص_2$ + $م^2$ $س_2^2$ − ٢ب $ص_2$ + ٢م ب $س_2$+$ب^2$

٣) $ص_3^2$ − ٢م $س_3$ $ص_3$ + $م^2$ $س_3^2$ ٢ب $ص_3$ + ٢م ب $س_3$ + $ب^2$.

وهكذا حتى نصل الى الفرق النوني وبالجمع نحصل على :

مج $ص^2$ − ٢م مج س ص + $م^2$ مج $س^2$ − ٢ب مج ص + ٢م ب مج س + ن $ب^2$.

وبالتعويض عن مج ص بما تساويه م مج س٢ + ب مج س وعن مج ص بما تساويه م مج س + ن ب نحصل على :

مج ص٢ − م مج س ص − ب مج ص − م٢ مج س٢ − م ب مج س − ن ب٢ + م٢ مج س٢ + ٢م ب مج س + ن ب٢ − م ب مج س .

وهذا يساوي بعد الاختصار مج ص٢ − م مج س ص − ب مج ص
أي مج ص٢ − (م مج س ص + ب مج ص)

وحيث ان مربع خطأ المعادلة يساوي متوسط هذه الفروق :

مربع خطأ المعادلة = $\frac{1}{ن}$ (مج ص٢) − (م مج س ص + ب مج ص)

وبتحليل القيمه $\frac{\text{مج ص}^2 - (\text{م مج س ص} + \text{ب مج ص})}{ن}$ أي

بالتعويض عن مج ص٢ بما تساويه من قانون الانحراف المعياري ع ص = $\frac{\text{مج ص}^2}{ن}$ − ص̄٢ وعن مج س ص بما تساويه من قانون الارتباط ر = $\frac{\text{مج س ص} \quad \text{ن س̄ ص̄}}{\text{ن ع}_س \text{ ع}_ص}$

وعن مج ص بما تساويه ن ص̄ وعن ب بما تساويه ص̄ − م س̄ .

∴ ح٢ =

$$\frac{\text{ن ع}^2_{ص} + \text{ن ص̄}^2 - \text{ن م ر ع}_{س} \text{ ع}_{ص} - \text{ن م س̄ ص̄} - \text{ن ص̄}^2 + \text{ن م س̄ ص̄}}{ن}$$

∴ ع٢ص − م ر ع س ع ص = مربع خطأ المعادلة

وبالقسمة على ع٢ ص نحصل على :

$$\text{١} - \text{م}\ \text{ر}\ \frac{\text{ع}_{\text{س}}}{\text{ع}_{\text{ص}}} = \frac{\text{مربع خطأ المعادلة}}{\text{ع}^{\text{٢}}_{\text{ص}}}$$

ولكن $\text{م} = \text{ر}\ \frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}}$

$$\therefore\ \text{١} - \text{ر}\ \frac{\text{ع}_{\text{ص}}}{\text{ع}_{\text{س}}} \times \text{ر}\ \frac{\text{ع}_{\text{س}}}{\text{ع}_{\text{ص}}} = \frac{\text{مربع خطأ المعادلة}}{\text{ع}^{\text{٢}}_{\text{ص}}}$$

$$\therefore\ \text{١} - \text{ر}^{\text{٢}} = \frac{\text{مربع خطأ المعادلة}}{\text{ع}^{\text{٢}}_{\text{ص}}}$$

$$\therefore\ \text{ر}^{\text{٢}} = \text{١} - \frac{\text{مربع خطأ المعادلة}}{\text{ع}^{\text{٢}}_{\text{ص}}}$$

أي $\text{ر}^{\text{٢}} = \text{١} - \frac{\text{خ}^{\text{٢}}_{\text{ص}}}{\text{ع}^{\text{٢}}_{\text{ص}}}$

دلالة معامل الارتباط :

ان اعتمادنا على معامل الارتباط لقياس درجة الارتباط بين ظاهرتين يمكن أن يضللنا حيث قد يحدث تلازم بين قيم الظاهرتين لمجرد الصدفة . ولا شك أن تأثير عامل الصدفة على النتيجة التي نصل اليها عند قياس الارتباط يزداد كثيراً كلما نقصت عدد الحالات التي أجرينا عليها البحث ، وبذلك تكون دلالة معامل الارتباط دلاله ضعيفة . فلو فرضنا اننا درسنا حالتين فقط (وحدتين فقط) فان الارتباط بينهما لا بد أن يكون كاملا مهما كانت قيم المتغيرين ومن الواضح أن هذه النتيجة قد تكون بعيدة كلياً عن الحقيقة — إلا انه كلما زاد عدد أزواج القيم المتناظرة التي حسب الارتباط على أساسها

فان احتمال الحصول على معامل ارتباط قوي بين الظاهرتين بطريق الصدقة يكون احتمال ضعيف ويزداد ضعف الاحـــتمال كلما زاد عدد أزواج قيم الظاهرتين – فاذا توصلنا إلى معامل ارتباط ٤ر٠ من دراسة ١٠٠٠ حالة تكون دلالة هذا المعامل أكبر بكثير من دلالة معامل ارتباط ٧ر٠ نتج عن دراسة عشر حالات فقط .

ولقد أعد الجدول الآتي ليبين احتمالات الحصول على قيم مختلفة لمعامل الارتباط بطريق الصدفة حسب عدد أزواج القيم المتناظرة :

عدد ازواج القيم المتناظـرة	احتمال الحصول على قيمة ر المبينة في هذا الجدول بطريق الصدفـة			
	١ر	٠٥ر	٠٢ر	٠١ر
١٠	ر = ٠ر٥٥	ر = ٠ر٦٣	ر = ٠ر٧٢	ر = ٠ر٧٦
١٥	٠ر٤٤	٠ر٥١	٠ر٥٩	٠ر٦٤
٢٠	٠ر٣٨	٠ر٤٤	٠ر٥٢	٠ر٥٦
٢٧	٠ر٣٢	٠ر٣٨	٠ر٤٥	٠ر٤٩
٣٢	٠ر٣٠	٠ر٣٥	٠ر٤١	٠ر٤٥
٣٧	٠ر٢٨	٠ر٣٢	٠ر٣٨	٠ر٤٢
٤٢	٠ر٢٦	٠ر٣٠	٠ر٣٦	٠ر٣٩
٥٠	٠ر٢٣	٠ر٢٨	٠ر٣٣	٠ر٣٥
٦٠	٠ر٢١	٠ر٢٥	٠ر٣٠	٠ر٣٣
٧٠	٠ر٢٠	٠ر٢٣	٠ر٢٨	٠ر٣٠
٨٠	٠ر١٨	٠ر٢٢	٠ر٢٦	٠ر٢٨
٩٠	٠ر١٧	٠ر٢٠	٠ر٢٤	٠ر٢٧
١٠٠	٠ر١٦	٠ر٢٠	٠ر٢٣	٠ر٢٦

ويمكن اختيار دلالة معامل الارتباط بطريقة اخرى ذلك أننا غالباً في قياس الارتباط نعتمد على عينة من المجتمع موضوع البحث ، وبذلك على فرض

عدم وجود ارتباط بين الظاهرتين موضوع البحث فان معامل الارتباط لو حسب للمجتمع كله وليس للعينة فقط كانت النتيجة تساوي صفراً . فاذا اخذنا عدداً كبيراً جداً من العينات من هذا المجتمع وحسبنا لكل منها معامل الارتباط ثم وزعنا هـذه المعاملات في توزيع تكراري لحصلنا على توزيع معتدل متوسطه أي ان معامل الارتبـاط المتوسط (الواقعي للمجتمع) = صفر . ومن الواضـح ان الانحراف المعياري لهذا التوزيع يدل على حجم الاخطاء (الفروق) التي يحتمل ان تظهر بين نتائج العينات والنتيجة الواقعية (صفر) ، وبعبارة اخرى يكون هذا الانحراف المعياري هو الخطأ المعياري لمعامل الارتباط .

$$\text{والخطأ المعياري لمعامل الارتباط} = \frac{١}{\sqrt{ن - ١}}$$

وتبعاً لخصائص التوزيع المعتدل يكون الفرق الذي يساوي ضعف هذا هذا الخطأ في الحد الأقصى فرق راجع الى الصدفة (بدرجة ثقة ٩٥ر٤ ٪) ، أما الفرق الذي يساوي أكثر من ضعفي الخطأ المعياري يكون فرقاً جوهرياً ، وبذلك اذا كان معامل الارتباط الذي نحصل عليه من العينة (يؤخـذ المعامل نفسه لأن فرقه عـن الصفر = ر) أقل من ضعف الخطأ المعياري نستطيع أن نستنتج ان هناك احتمال كبير (٩٥ر٤ ٪) في أن يكون هذا المعامل راجعاً الى الصدفة ، واذا كان معامل الارتباط اكبر من ضعف الخطأ المعياري نستطيع أن نستنتج ان هنـاك احتمال صغير (١٠٠ - ٩٥ر٤ = ٤ر٦ ٪) في أن يكون هذا المعامل راجعاً الى الصدفة بمعنى أن هـذا المعامل له دلالته ومعنويته في قياس الارتباط بين الظاهرتين :

مثال ٦٥ :

فاذا حصلنا على معامل ارتباط ٧ ر من دراسة ١٠٠ حالة

يكون الخطأ المعياري لمعامل الارتباط = $\frac{1}{\sqrt{\text{ن} - 1}} = \frac{1}{\sqrt{99}}$

= ٥ ١٠٠ ر

∴ ٢ ع = ٢ × ١٠٠٥و. = ٢٠١٠و.

وحيث ان معامل الارتباط ٧ر أكبر بكثير من هذا الرقم ، فلا يمكن بأي حال من الأحوال ان يكون هذا المعامل راجعاً الى الصدفة ، أي ان هذا المعامل يدل دلالة حقيقية على وجود الارتباط بين الظاهرتين موضوع الدرس .

تمارين

١) من القيم الآتية للمتغيرين س ،ص أوجد معادلتي الانحدار بطريقتين :

س	٩ر٢	١٠ر٦	١٠ر٣	١٣ر٨	١٧ر٥	٢٠ر١	١٦ر٩	١٤ر٦	١٧	١٦
ص	١٠ر٥	٩ر٢	٩ر٧	١١ر٥	١٢ر٥	١٣	١٤ر٣	١٥	١٣ر٩	١٧ر٤

(ص = ٤٦٥ر٠ س + ٩ر٥)

(س = ٩١٤ر٠ ص+ ٣)

٢) احسب معامل الارتباط بين س ، ص بطريقة سبيرمان :

س	١٠	١٢	١٣	١٦	١٧	٢٠	١٦	١١
ص	١٩	٢٢	٢٣	٢٤	٢٢	٢٧	٢٢	٣٣

٣) من الجدول التالي اوجد معامل الارتباط بين الظاهرتين س ، ص :

ص	س				المجموع
	٢ ـ	٤ ـ	٦ ـ	٨ ـ	
١٠ ـ	٢	٤	ـ	ـ	٦
٢٠ ـ	١	٦	٢	ـ	٩
٣٠ ـ	١	٤	٤	ـ	٩
٤٠ ـ	ـ	٣	٨	٣	١٤
٥٠ ـ ٦٠	ـ	ـ	٥	٧	١٢
المجموع	٤	١٧	١٩	١٠	٥٠

ثم اوجد معادلتي الانحدار ـ اختبر دلالة معامل الارتباط .

٤ ـ الجدول الآتي يبين توزيع ١٠٠ أسرة تبعاً لعدد حجرات السكن (ص) والدخل الشهري (س) .

س	ص					المجموع
	١	٢	٣	٤	٥	
٥ ـ	٥	٥	٣	١	ـ	١٤
١٠ ـ	٣	٧	١٠	٤	ـ	٢٤
١٥ ـ	٢	٩	٢٢	٥	ـ	٣٨
٢٠ ـ	ـ	٤	٨	٣	١	١٦
٢٥ ـ ٣٠	ـ	٢	٢	ـ	٤	٦
المجموع	١٠	٢٧	٤٥	١٣	٥	١٠٠

المطلوب :

١ – رسم شكل انتشار يبين انحدار الايجار على عدد الغرف والتعليق عليـــه

٢ – حساب معامل الارتباط واختبار دلالته .

٣ – ايجاد معادلة انحدار الايجار على عدد الغرف .

٤ – حساب خطأ المعادلة – فسر النتيجة .

٥ – تقدير متوسط الدخل الذي يدفع لأربع غرف وتحقيق النتيجة من الجدول .

الفصل الثالث عشر

الارقام القياسية

نشأ الاهتمام بموضوع الارقام القياسية نتيجة الحاجة إلى مقياس محدد يقيس التغيرات في قيمة النقود. على ان ذلك ليس هو المجال الوحيد لاستخدام الارقام القياسية ، إذ قد اتسع مجال استخدامها حتى أصبح يشمل كثيراً من الظواهر الاقتصادية وغير الاقتصادية .

والرقم القياسي هو مقياس احصائي لقياس التغيرات التي تحدث في متغير واحد أو في مجموعة من المتغيرات خلال الزمن أو من منطقة جغرافية إلى اخرى أو على اساس اختلاف الدخول والمهن ... الخ .

وباستعمال الارقام القياسية نستطيع مثلا مقارنة تكلفة الطعام أو أي بند آخر من بنود الانفاق هذا العام بالنسبة للعام السابق، أو نستطيع مقارنة انتاج الصلب في عام معين في منطقة معينة من الدولة بالنسبة لمنطقة اخرى في نفس العام . وبالرغم من ان الارقام القياسية تستخدم اساساً في قياس التغيرات الخاصة بالظواهر الاقتصادية ، الا انه يمكن تطبيقها كذلك في نواحي اخرى كثيرة ، مثلا في مجال التربية نستطيع باستخدام الارقام القياسية في مقارنة ذكاء الطلاب في مناطق الدولة او في سنوات مختلفة .

وتهتم كثير من الادارات الحكومية والخاصة بتركيب أرقام قياسية يمكن استعمالها في توقع المستقبل بالنسبة للاعمال التجارية والنشاطات

الاقتصادية المختلفة ، ولذلك نجد في كثير من الدول أرقاماً قياسية للاجور واخرى للانتاج بأنواعه المختلفة واخرى للبطالة ... الخ . ولعل أكثر هذه الارقام شيوعاً هو الرقم القياسي لنفقة المعيشة أو كما يسمى أحياناً الرقم القياسي لاسعار السلع الاستهلاكية ، وفي بعض الدول تحتوي عقود العمل الجماعية على نص بزيادة الاجور تلقائياً كلما دلت هذه الارقام القياسية على الارتفاع .

ويمكن أن نعطي هنا أمثلة لبعض التطبيقات العملية للارقام القياسية :

١ – يقاس التغير في الأسعار في فترة زمنية معينة وذلك لاكتشاف سبب التغير وأثره على النشاط الاقتصادي حتى يمكن التحكم فيه .

٢ – قياس التغير في نفقة المعيشة وخاصة بالنسبة للطبقة العاملة حتى يمكن البت في مطالبهم الخاصة بالاجور .

٣ – قياس التغير في مستوى معيشة طبقة معينة من الناس ، وذلك بقسمة الرقم القياسي لدخولهم النقدية على الرقم القياسي لنفقة معيشتهم فنحصل على رقم قياسي للأجر الحقيقي يبين لنا التغير في مستوى معيشة هذه الطبقة فنستطيع أن نقرر مساعدتهم إذا تبين لنا أن التغير في حالتهم يدعو إلى ذلك .

٤ – قياس التغير في حجم الانتاج الصناعي والزراعي والموجودات في المخازن والتجارة الخارجية والمبيعات .

٥ – قياس التغير في عدد المشتغلين عامة وفي كل ناحية من نواحي النشاط الاقتصادي على حدة ، وكذلك التغير في عدد المتعطلين .

تكوين الرقم القياسي

يواجه الاحصائي المشاكل الآتية عند تكوين رقم قياسي :

١ – اختيار المتغيرات التي يشملها الرقم .

٢ – اختيار المصادر التي يجمع منها البيانات اللازمة .

٣ – اختيار الفترة أو المكان الذي يعتبر أساساً لقياس التغير

٤ – اختيار الصيغة التي تستحدم في حساب الرقم القياسي حيث توجد عدة صيغ سيأتي شرحها فيما بعد .

٥ – اختيار الأوزان التي تستخدم في الترجيح إذا اتفقنا على صيغة يكون فيها الرقم القياسي مرجحاً .

٦ – تنظيم العمل بحيث يمكن حساب الرقم سريعاً حتى تتحقق فائدته العملية إذ لا فـــائدة عملية لرقم قياسي ينشر بعد فوات الوقت المتعلق به .

ولشرح هذه المشاكل سنقتصر في دراستنا على الأرقام القياسية للاسعار مع ملاحظة أن ما ينطبق على هذا النوع من الأرقام القياسية ينطبق أيضاً على الأنواع الاخرى مع بعض التعديلات الطفيفة التي تناسب كل نوع .

المشكلة الاولى :

عند تكوين رقم قياس للاسعار يجب أن نحدد انواع السلع التي يشملها الرقم ، وبالطبع يتوقف هذا التحديد على نوع الرقم ؛ فـاذا كنا نبغي رقماً يقيس التغير في نفقة المعيشة لطبقة معينه من الناس يجب أن نقتصر على السلع التي تستهلكها هذه الطبقة ، وإذا كنا نبغي رقماً يقيس التغير في نفقات الانتاج الصناعي اقتصرنا على السلع التي تعتبر مـــواداً أولية في الصناعة .

وبذلك يكون أول ما يجب أن نراعيه عند تكوين رقماً قياسياً أن نختار عينة ممثلة تمثيلاً صادقاً للموضوع الذي نريد دراسته ، ولكي تكون العينة

ممثلة يجب أن تشمل عدداً لا بأس به من السلع ، وليس هناك قاعدة محدودة يمكن أن نراعيها عند تحديد عدد السلع وان كان يجب أن نلاحظ أن لا يزيد العدد كثيراً لأن في ذلك تعقيد للعمليات الحسابية ، الامر الذي يؤخر ظهور النتائج ، والا يقل العدد كثيراً حيث تصبح العينة نتيجة لذلك غير ممثلة تمثيلا صادقاً .

ويحسن تقسيم السلع التي تدخل في الرقم إلى مجموعات ، تتميز المفردات في كل منها بصفات خاصة ذات اهمية في الناحية التي يتناولها الرقم القياسي الذي نريد تكوينه . فمثلا عند تكوين رقم قياسي لنفقة المعيشة نقسم السلع إلى مجموعة اللحوم ، والخضروات ، والفواكه ، والملابس .. الخ ولا يجب أن نأخذ جميع السلع التي تتكون منها كل مجموعة فيكفي أن نأخذ بعضها والتي تكون شائعة الاستعمال لدى الطبقة التي يراد قياس نفقة معيشتها والتي تمثل الاتجاهات المختلفة في الاسعار بمعنى انه إذا كان هناك عدة سلع تسير اسعارها سوياً في اتجاه واحد سواء بالارتفاع أو الانخفاض يكفي أن نأخذ منها واحدة أو اثنتين ، أما السلع التي تتغير اسعارها في اتجاهات مختلفة فيحسن أن نأخذها كلها .

المشكلة الثانية :

عند اختيار المصادر التي تجمع منها البيانات الخاصة بالرقم القياسي يجب أن نلجأ إلى التجار الذين نعتقد ان افراد الطبقة التي نريد قياس نفقة معيشتها يشترون منهم حاجاتهم . كذلك يجب أن نتأكد من أن الأسعار التي نجمعها هي عن السلع التي قررنا إدخالها في الرقم مع ملاحظة الانواع المختلفة من السلعة الواحدة ، فاذا أردنا أن نجمع بيانات عن أسعار السيارات فنلاحظ أنواعها المختلفة وكذلك ان النوع الواحد يختلف سعره حسب السنة التي أنتج فيها ، فسيارة موديل ١٩٦٥ يختلف سعرها عن سيارةموديل ١٩٤٠ . كذلك

يجب أن نراعي تغير استعمال السلع من وقت الى آخر ، فسلعة ما قد تكون شائعة الاستعمال في وقت معين ولكن يبطل استعمالها في وقت آخر ، الأمر الذي يجعلنا نعمل دائماً على تعديل الرقم القياسي من وقت الى آخر (جميع الدول عـــدلت الرقم القياسي لنفقة المعيشة فيها حتى يتفق مع تغير طرق المعيشة) .

كذلك عند جمع البيانات عن الاسعار يجب ملاحظة الآنية في ذاك بمعنى أن نجمع الاسعار في نفس الوقت ولكي ندرك أهمية ذلك نتذكر التغييرات التي تصيب الاسعار من وقت إلى آخر وعند تحديد الوقت يجب مراعاة الفترة التي يقبل الأفراد فيها فعلاً على الشراء .

ومراعاة الدقة في جمع البيانات عن الأسعار من الأهمية بمكان كبير إذ ان أي خطـــأ لا بد أن يظهر تأثيره فى النتيجة النهائية للرقم القياسي وذلك بعكس الخطأ في الاوزان التي ترجح بها فإنه لا يؤثر كثيراً في النتيجة، وذلك لأن الوزن المضروب في السعر أو فى منسوبه يوجـــد في البسط وفي المقام أيضاً . أما إذا أخطأنا في السعر فإن البسط يتغير وحده دون المقام فيكون التغيز في قيمة الكسر أكثر منه في الحالة الأولى . وعلى ذلك يجب ان نعنى عناية تامة عند جمع بيانات الاسعار .

المشكلة الثالثة :

لتركيب الرقم القياسي لأية ظاهرة تكوّن نسبة مئوية بين القيمة المقارنة لهذه الظاهرة أو قيمتها في الفترة التي نريد مقارنتها والقيمة الاخرى المعتبرة أساساً للمقارنة ، فالرقم القياسي لسعر القمح في سنة ١٩٦٥ مثلا هو عبارة النسبة المئوية لسعر القمح في سنة ١٩٦٥ الى سعره في سنة اخرى نتفق على اعتبارها أساساً للمقارنة .

وبذلك عند إنشاء أي رقم قياسي يجب أن نحدد الاساس الذي سنتخذه

لترتيب الرقم ، فإذا كانت المقارنة على أساس الزمن يجب أن نحدد الفترة التي نعتبرها أساساً ، وإذا كانت المقارنة على أساس المكان يجب أن نحدد المكان الذي نعتبره أساساً .

وعند اختيار السنة الأساس نراعي ان تكون السنة تتميز بالاستقرار في الاسعار بمعنى أن لا تكون سنة شاذة من ناحية النشاط الاقتصادي فلا تكون سنة كساد أو سنة رواج . ولعل هذا هو السبب في أن معظم الدول كانت قبل الحرب العالمية الثانية تستخدم سنة ١٩١٣ أساساً للأرقام القياسية التي تكوّنها . ثم أصبحت تستخدم سنة ١٩٣٨ أساساً بعد الحرب العالمية الثانية حيث ان عام ١٩١٣ أصبح بعيداً جداً والمقارنة معه تكون مضللة .

ولكن بالرغم من أن سنة ما قد تكون مناسبة كأساس في مرحلة تاريخية معينة فإنها تصبح غير مناسبة بمرور الزمن ولهذا يحسن تغيير فترة الأساس إلى فترة قريبة من الفترة التي نقارنها . ويرجع ذلك الى عدة أسباب منها تغير طرق الاستهلاك وتغير أنواع السلع .

ويتبع أحياناً طريقة السلسلة في تكوين الرقم القياسي بمنى أن يحسب الرقم على أساس متحرك فلا تكون هناك فترة ثابتة تقارن الاسعار على أساسها ، وانما تكون فترة الاساس متغيرة حيث تكون سلسلة من الارقام القياسية كل سنة بالنسبة الى سابقتها وذلك إذا لم نضمن ثبات الظروف الاقتصادية ، وبذلك يصبح الرقم القياسي مرناً مرونة لا تتوفر في الرقم القياسي للقاعدة الثابتة حيث لا يتمشى في هذه الحالة مع الظروف الاقتصادية التي تحيط بالسلع الداخله في تركيب الرقم .

المشكلة الرابعة :

إذا قسمنا سعر أي سلعة في سنة ما على سعرها في سنة الاساس وضربنا الناتج في ١٠٠ فإننا نحصل على ما يسمى بمنسوب السعر ، وهو اصطلاح نستخدمه في في حالة تكوين رقم قياسي لسلعة واحدة فقط . أما إذا أدخلنا في تكوين

الرقم أكثر من سلعة واحدة فإننا نسمي الرقم الناتج بالرقم القياسي

ولتكوين رقم قياسي هناك طريقتان: طريقة التجميع وطريقة المناسيب. وهناك صيغ مختلفة لكل من الطريقتين

طريقة التجميع – الرقم التجميعي البسيط للاسعار :

لتركيب هذا الرقم يقسم مجموع أسعار السلع في سنة المقارنة على مجموع أسعارها في سنة الأساس ونضرب الناتج في ١٠٠ :

أي $\frac{مجـ ع_١}{مجـ ع.} \times ١٠٠$ — ١ —

مثال ٦٦ :

السلع	أ	ب	جـ	د
الأسعار في سنة ١٩٣٠	٢٠	١٢	٢٢	٩
الأسعار في سنة ١٩٥٠	٤١	١٩	٢٩	١١

الرقم التجميعي البسيط $= \frac{٤١ + ١٩ + ٢٩ + ١١}{٢٠ + ١٢ + ٢٢ + ٩} \times ١٠٠$

وميزة هذا الرقم هي البساطة في حسابه ولكن هذه الميزة هي عيب يؤخذ عليه في نفس الوقت ذلك لأن في استعمال هذه الصيغة نعتبر جميع السلع متعادلة في الأهمية من حيث تأثيرها على المتوسط العام للأسعار وهذا اعتبار خاطىء فلكل سلعة أهميتها وتأثيرها الذي يتناسب مع تلك الأهمية .

الارقام التجمعية المرجحة

لتصحيح الرقم التجمعي البسيط ترجح أسعار السلع المختلفة بأوزان

تتناسب وأهمية هذه السلع فتستخدم الكميات المنتجة أو المستهلكة من هذه السلع كأوزان ، ويصح أن نستعمل كميات سنة المقارنة أو سنة الأساس أو متوسطها . وهناك أنواع مختلفة من الأوزان سيأتي شرحها فيما بعد والصيغ المختلفه للأرقام المرجحة هي :

١ – الرقم التجمعي المرجح بكميات سنة الأساس
(رقم لاسبير)

$$\frac{\text{مجـ}\ \text{ع}_{١}\,\text{ك}_{٠}}{\text{مجـ}\ \text{ع}_{٠}\,\text{ك}_{٠}} \times ١٠٠$$

٢ الرقم التجمعي المرجح بكميات سنة المقارنة
(رقم باشه)

$$\frac{\text{مجـ}\ \text{ع}_{١}\,\text{ك}_{١}}{\text{مجـ}\ \text{ع}_{٠}\,\text{ك}_{١}} \times ١٠٠$$

٣ – الرقم التجمعي المرجح بالوسط الحسابي لكميات الأساس والمقارنـــة .
(الرقم القياسي لمارشال وادجورث)

$$\frac{\text{مجـ}\ \text{ع}_{١}\,(\text{ك}_{٠} + \text{ك}_{١})}{\text{مجـ}\ \text{ع}_{٠}\,(\text{ك}_{٠} + \text{ك}_{١})} \times ١٠٠$$

٤ – الرقم التجمعي المرجح بالوسط الهندسي لكميات الأساس والمقارنـــة .

$$\frac{\text{مجـ}\ \text{ع}_{١}\sqrt{\text{ك}_{٠} \times \text{ك}_{١}}}{\text{مجـ}\ \text{ع}_{٠}\sqrt{\text{ك}_{٠} \times \text{ك}_{١}}} \times ١٠٠$$

٥ الرقم الأمثل وهو عبارة عـن الوسط الهندسي للرقمـــين الأول والثاني .
(رقم فيشر)

$$\sqrt{\frac{\text{مجـ}\ \text{ع}_{١}\,\text{ك}_{٠}}{\text{مجـ}\ \text{ع}_{٠}\,\text{ك}_{٠}} \times \frac{\text{مجـ}\ \text{ع}_{١}\,\text{ك}_{١}}{\text{مجـ}\ \text{ع}_{٠}\,\text{ك}_{١}}} \times ١٠٠$$

مج ع١ ك٠ / مج ع٠ ك٠ × ١٠٠

مثال ٦٧ :

السلعة	كمية سنة الاساس	كمية سنة المقارنة	أسعار سنة الاساس	أسعار سنة المقارنة
أ	٢٥٥	٢٨٠	٢٠	٤١
ب	١٨٥	٢٢٥	٢٢	١٩
حـ	٣٠٥	٤١٨	٢٢	٢٩
د	١١٠	١٢٤	٩	١١

الرقم / ١ =

٤١ × ٢٥٥ + ١٩ × ١٨٥ + ٢٩ × ٣٠٥ + ١١ × ١١٠
―――――――――――――――――――――――――― × ١٠٠
٢٠ × ٢٥٥ + ١٢ × ١٨٥ + ٢٢ × ٣٠٥ + ٩ × ١١٠

الرقم / ٢ =

٤١ × ٢٨٠ + ١٩ × ٢٣٥ + ٢٩ × ٤١٨ + ١١ × ١٢٥
―――――――――――――――――――――――――― × ١٠٠
٢٠ × ٢٨٠ + ١٢ × ٢٣٥ + ٢٢ × ٤١٨ + ٩ × ١٢٥

الرقم / ٣ =

٤١×(٢٥٥+٢٨٠)+١٩(١٨٥+٢٣٥)+ ٢٩(٣٠٥+٤١٨)+١١(١٠٠+١٢٥)
――――――――――――――――――――――――――――――
٢٠(٢٥٥+٢٨٠)+١٢(١٨٥+٢٣٥) +٢٢ (٣٩٥+٤١٨) +٩ (١١٠+١٢٥)
× ١٠٠

الرقم / ٤ =

٤١ ٢٥٥×٢٨٠+١٩ ١٨٥×٢٣٥+٢٩ ٣٠٥×٤١٨+١١ ١١٠×١٢٥
――――――――――――――――――――――――――――――
٢٠ ٢٥٥×٢٨٩+١٢ ١٨٥×٢٣٥+٢٢ ٣٠٥×٤١٨+٩ ١١٠×١٢٠
× ١٠٠

الرقم / ٥ يمكن حسابه بايجـــاد الوسط الهندسي لنتيجة الرقم الاول والثـــاني .

طريقة المناسيب :

يمكن حساب الرقم القياسي على أساس الوسط لمناسيب اسعار السلعة الداخلة في تكوين الرقم ، والوسط امـــا أن يكون حسابياً او هندسياً أو توافقياً ، وهذه المتوسطات اما ان تكون بسيطة او مرجحة .

ولحساب الرقم القياسي تبعاً لهذه الطريقة نبدأ بحساب منسوب السعر لكل سلعة تدخل في تكوين الرقم .

منسوب سعر السلعة أ $= \frac{ع_1}{ع_0} \times 100 = س_1$

منسوب سعر السلعة ب $= \frac{ع_1}{ع_0} \times 100 = س_2$

منسوب سعر السلعة جـ $= \frac{ع_1}{ع_0} \times 100 = س_3$

منسوب سعر السلعة د $= \frac{ع_1}{ع_0} \times 100 = س_4$

وبذلك يكون الرقم القياسي لأسعار هذه السلع باستخدام الوسط الحسابي البسيط وهو :

$$\frac{س_1 + س_2 + س_3 + س_4}{4}$$

والرقم القياسي لأسعار هذه السلع باستخدام الوسط الهندسي البسيط هو :

$$\sqrt[4]{س_1 \times س_2 \times س_3 \times س_4}$$

والرقم القياسي لاسعار هذه السلع باستخدام الوسط التوافقي البسيط هو :

$$\frac{4}{\frac{1}{س_1} + \frac{1}{س_2} + \frac{1}{س_3} + \frac{1}{س_4}}$$

هذه المتوسطات البسيطة للمناسب لا تفرق بين السلع المختلفة بل تعاملها جميعاً نفس المعاملة وهذا عيب الارقام القياسية التي تكون تبعاً لهذه الطريقة ، ذلك لان السلع تختلف من حيث أهميتها وتأثير ذلك على المتوسط العام للأسعار . لذلك يحسن تكوين الارقام القياسية بترجيح مناسيب أسعار السلع تبعاً لاهميتها . وللترجيح يمكن استخدام قيم السلع كأوزان . هذه القيم يمكن أن تحسب تبعاً للتوافيق الآتية :

١ - سعر سنة الاساس × كمية سنة الاساس = $ع_0$ $ك_0$ = $م_0$ مثلاً

٢ - سعر سنة المقارنة × كمية سنة المقارنة = $ع_1$ $ك_1$ = $م_1$ مثلا

٣ - سعر سنة الاساس × كمية سنة المقارنة = $ع_0$ $ك_1$ = $م_2$ مثلا

٤ - سعر سنة المقارنة × كمية سنة الاساس = $ع_1$ $ك_0$ = $م_3$ مثلا

وبذلك يمكن استخدام احدى الصيغ الآتية لحساب الرقم القياسي تبعاً للوسط الحسابي المرجح :

١) $$\frac{\sum س\ م_0}{\sum م_0}$$

٢ — $\frac{\text{مجـ س م}_{١}}{\text{مجـ م}_{١}}$

٣ — $\frac{\text{مجـ س م}_{٢}}{\text{مجـ م}_{٢}}$

٤ — $\frac{\text{مجـ س م}_{٣}}{\text{مجـ م}_{٣}}$

وبالمثل نحصل على الوسط التوافقي المرجح بهذه الاوزان الاربعة كالآتي :

١ — $\frac{\text{مجـ م}_{٠}}{\text{مجـ}\left(\frac{\text{م}_{٠}}{\text{س}}\right)}$

٢ — $\frac{\text{مجـ م}_{١}}{\text{مجـ}\left(\frac{\text{م}_{١}}{\text{س}}\right)}$

٣ — $\frac{\text{مجـ م}_{٢}}{\text{مجـ}\left(\frac{\text{م}_{٢}}{\text{س}}\right)}$

٤ — $\frac{\text{مجـ م}_{٣}}{\text{مجـ}\left(\frac{\text{م}_{٣}}{\text{س}}\right)}$

كذلك يمكن أن نحصل على الوسط الهندسي المرجح بهذه الاوزان الاربعة كالآتي :

$$\text{١ – مح م}_{.}$$
$$\sqrt{\text{(س)م}_{.} \times \text{(س)م}_{.} \times \text{(س)م}_{.}}$$

$$\text{٢ – مح م}_{١}$$
$$\sqrt{\text{(س)م}_{١} \times \text{(س)م}_{١} \times \text{(ش)م}_{١}}$$

$$\text{٣ – مح م}_{٢}$$
$$\sqrt{\text{(س)م}_{٢} \times \text{(س) م}_{٢} \times \text{(س)م}_{٢}}$$

$$\text{٣ – مح م}_{٤}$$
$$\sqrt{\text{(س)م}_{٤} \times \text{(س)م}_{٤} \times \text{(س)م}_{٤}}$$

طريقة السلسلة :

في تكوين الارقام القياسية تبعاً للطرق السابقة كنا ننسب الاسعار في سنة المقارنة الى اسعار سنة أساس ثابتة – ولكن اذا طالت المدة بين سنة الاساس وسنة المقارنة فإن الزمن كفيل بأن يحدث تغييراً كبيراً في الظروف المحيطة بالسلع التي نبحثها والتي يتركب منها الرقم القياسي . فقد يحدث أن بعض السلع الشائعة الاستهلاك تنعدم في الأزمنة الحديثة وبالعكس قد توجد بعض السلع لم تكن معروفة من قبل ، أو على الأقل تتغير أهمية السلع . فإذا حدث ذلك فإننا لا نستطيع أن نطمئن الى صحة المقارنة ، ولتلافي ذلك يحسن تكوين الارقام القياسية كل سنة بالنسبة الى سابقتها كأساس .

ولتكوين الرقم القياسي تبعاً لهذه الطريقة ننسب الاسعار في سنة ١٩٦٤ مثلاً الى الاسعار في سنة ١٩٦٣ ، والاسعار في سنة ١٩٦٥ إلى أسعار ١٩٦٤، والاسعار في سنة ١٩٦٦ إلى الاسعار في ١٩٦٥ مستخدمين أي معادلة من المعادلات التي سبق ذكرها ؛ وبذلك يكون الاساس في هذه السلسلة متحركا وليس ثابتاً كما لو اتخذنا سنة ١٩٥٣ أساساً لكل السنين .

وإذا أردنا مقارنة الاسعار في فترة معينة تبعد عنها بضع فترات وكان لدينا الارقام القياسية المحسوبة على الاساس المتحرك فما علينـــا إلا ضرب الارقام المتتالية حتى نصل إلى الفترة المطلوبة ثم قسمة حـــاصل الضرب على $١٠٠^{ن-١}$ حيث ن عدد الارقام المضروبة في بعضها البعض .

مثال ٦٨ :

إذا رمزنا للرقم القياسي لسنة ١٩٦٤ بالنسبة لسنة ١٩٦٣ بالرمز $ا_١$

» » » » ١٩٦٥ » » ١٩٦٤ » $ا_٢$

» » » » ١٩٦٦ » » ١٩٦٥ » $ا_٣$

ثم أردنا ايجاد الرقم القياسي في سنة ١٩٦٦ بالنسبة لسنة ١٩٦٣ يكون كالآتي :

$$\frac{ا_١ \times أ_٢ \times ا_٣}{١٠٠ \times ١٠٠}$$

ونلاحظ ان النتائج التي نحصل عليها للرقم القياسي إذا استخدمنا أساساً ثابتاً أو أساساً متحركا تختلف عن بعضها ، كذلك إذا أرجعنا رقما قياسياً على أساس متحرك إلى أساس ثابت فان النتيجة تختلف عما لو حسب بالرقم القياسي على الأساس الثابت .

المشكلة الخامسة :

يمكن أن يكون الترجيح مباشراً أو غير مباشر . ففي حالة الترجيـــح المباشر تعطى كل سلعة داخلة في تكوين الرقم القياسي وزن معيّن يدل على أهميتها النسبية أي على أهميتها بالنسبة للسلع الأخرى التي تدخل في تكوين الرقم . ويتوقف اختيار الأوزان على البيانات الموجودة لدينا ، فإذا كان لدينا بيانات عن كميات الاستهلاك او الانتاج من السلع المختلفة في سنة الاساس فقط

اضطررنا إلى استخدام هذه الكميات بالرغم من أن سنة الاساس قد تكون بعيدة نوعاً ما ، الأمر الذى يجعل الاهمية النسبية للسلع المختلفة قد تغيرت في الفترة الطويلة بين سنة الاساس وسنة المقارنة . وإذا وجد لدينا بيانات عن هذه الكميات في سنة المقارنة فيكون من الافضل استخدامها كأوزان حيث انها تعبر تعبيراً دقيقـــاً عن الأهمية الحاضرة للسلع المختلفة ، وإن كنا نشك كثيراً في إمكان وجود بيانات عن هذه الكميات في الوقت الذي نريـــد فيه تكوين الرقم القياسي حيث ان الغالب ان هذا النوع من الكميات لا ينشر بيان عنها إلا بعد الفترة التي تتعلق بها .

وبالنسبة للرقم القياسي لنفقة المعيشة فإن الاوزان تحسب على أساس القيام ببحث استقصائي عن ميزانية الأسرة نتوصل منه الى معرفة متوسط نسبة ما ينفق من ميزانية الاسرة على السلع المختلفة التي تدخل في معيشتها .

أما الترجيح غير المباشر يكون بجعل عدد السلع والانواع التي تدخل في كل مجموعـــة من المجموعات التي يتكون منها الرقم القياسي يتناسب مع أهمية المجموعة ، ذلك لأن كثرة عدد السلع المأخوذة من أي مجموعة تعمل على تعزيز هذه المجموعة وترجيح التغييرات التي تحصل في أسعار سلعها وبالعكس فإن قلة عدد السلع المأخوذ في أي مجموعة ينقص من أهمية التغيرات التي تحصل في أسعار سلعها . كذلك يمكن أن يكون الترجيح الغير مباشر على أساس السلع وليس على أساس المجموعات وذلك بأن نأخذ عدة انواع لنفس السلعة وندخل أسعار هذه الانواع في تكوين الرقم القياسي وكأننا بذلك رجحنا هذه السلعة ترجيحاً غير مباشر وعززنا أهميتها بالنسبة الى غيرها .

وكما أشرنا سابقـــاً لا نكون في حاجة الى الدقة المتناهية في حساب الأوزان ، ذلك لأنها تضرب في البسيط وفي المقام وبذلـــك لا يؤثر الخطأ البسيط فيها على النتيجة تأثيراً كبيراً . ولهذا يحسن دائماً أن نأخذ أعداد صحيحة ونتغاضى عن الكسور حتى لا نعقد كثيراً من العمليات الحسابية .

المشكلة السادسة :

عند وضع خطة العمل لتركيب رقم قياسي يجب أن نلاحظ اننا نحتاج الى نشر النتيجة في الوقت المناسب كي يمكن الاستفادة منها فائدة عملية فلاتقتصر فائدتها للذكرى والتاريخ فقط ، وإذا تذكرنا ان بعض الارقام القياسية تنشر كل شهر أو ثلاثة أشهر أدركنا أهمية وصع الخطة المناسبة لتحقيق هـــذه الغاية ، ولهذا يحسن أن تكون هناك جداول معينة مدوّن بها كل البيانات المطلوبة عدا البيانات الخاصة بالأسعار في فترة المقارنة التي يجب أن يوكل جمعها الى موظفين مدربين بحيث يدركون أهمية جمع هذه البيانات في نفس الوقت وبدقة تامة وعن نفس السلع المحددة ، وأحب أن أنبه الى ان الاعتماد على النشرات الرسمية للأسعار لا يعطينا فكرة دقيقة إلا إذا تأكدنا ان التسعيرة الرسمية يتبعها التجار بدقة تامة ، ولكن الغالب ان تختلف الاسعار الواقعية اختلافاً كبيراً من الاسعار المحددة من قبل الحكومة ، ولهذا يحسن عدم الاعتماد عليها وأن تقوم الادارة الاحصائيـــة المختصة بجمع الاسعار التي يجري التعامل على أساسها في الاسواق المختلفة خاصة اذا تذكرنا ان الاسفار الرسمية تكون غالباً محددة لجميع أنواع الأسواق على السواء ، ولكن الواقع انالاسعار تختلفتبعاً لما إذا كان البائع في سوقالقرية أو فيسوق المدينة وتبعاً لما إذا كان البائع عاديا أو مؤسسة تعاونية وكذلك تبعـــاً للحي الذي يوجد به البائع , الوقت الذي تجمع فيه الاسعار .

ويلاحظ البعض ان قيام موظفو الادارة الاحصائية بجمع الاسعار قـــد يؤدي إلى عدم دقة البيانات الخاصة بها ، خاصة إذا كان هناك نظام للتسعيرة الجبرية تقوم الحكومة بفرضه على الاسواق . ولهذا يرون أن يوكل جمع الاسعار الى لجنة من ربات المنازل غير معروفات لدى التجار حتى تستطيـــع الادارة الاحصائية أن تطمئن إلى البيانات المجموعة .

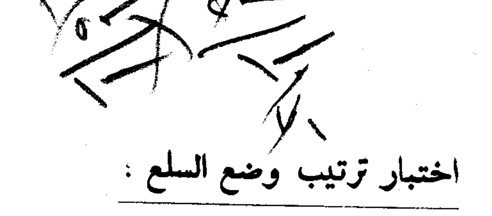
وهناك اعتبار آخر يجب ملاحظته ، ذلك اننا لا يجب أن نتمسك بصيغة

معينة للرقم القياسي لأننا نعتقد انها أفضل من ناحية الدقة ، إذ يجب أن نأخذ في اعتبارنا امكانية الحصول على البيانات اللازمة لحساب الرقم القياسي تبعاً لهذه الصيغة ، وهذا غالباً يتوقف على تقدم الوعي الاحصائي في الدولة عامة ولدى التجار على الوجه الاخص . ففي مصر مثلاً من السهولة أن تجمع بيانات عن الاسعار ولكن من الصعب معرفة بيانات عن الكميات المباعة أو المنتجة ، ذلك لأن التاجر نفسه لا يعرف ذلك إلا بعد الرجوع الى دفاتره إذا كانت لديه دفاتر منتظمة . هذا فضلاً عن ان هذا البيان بالذات يعتبره التاجر سراً من أسراره وليس عنده الاستعداد للادلاء به للغير . في مثل هذه الاحوال يجب استبعاد جميع الصيغ التي تدخل فيها ك لأن تمسكنا بواحدة من هذه الصيغ سيضيع وقتاً ، فضلاً عن اننا نعرف تماماً ان البيانات التي نجمعها ليست دقيقة .

اختبار الارقام القياسية :

رأينا كيف ان هناك صيغ مختلفة لتركيب الرقم القياسي والمفاضلة بينها لا بد ان تتناول الناحيتين العملية والنظرية ، وقد أشرنا من قبل الى المفاضلة على أساس الناحية العملية حيث رأينا كيف ان توفر البيانات والحاجة الى نشر النتائج في الوقت المناسب تحدد لنا الى حد بعيد الصيغة التي يحسن بنا استخدامها . أما المفاضلة على أساس الناحية النظرية فقد بحثها الاستاذ ارفتج فيشر حيث يرجع لديه الفضل الأكبر في هذا الموضوع .

اقترح فيشر عدة اختبارات بحيث إذا اجتازت احدى الصيغ هذه الاختبارات جميعها أمكن القول انها أفضل الصيغ . وقد وجد فيشر أن الصيغة التي تقوم على اساس الوسط الهندسي للرقم القياسي المرجح بكميات سنة الاساس والرقم المرجح بكميات سنة المقارنة هي أفضل الصيغ جميعاً حيث اجتازت الاختبارات التي يقترحها ولذا أطلق عليه الرقم القياسي الامثل .



اختبار ترتيب وضع السلع :

من الواضح ان النتيجة التي نحصل عليها من استخدام أي صيغة للرقم القياسي لا يجب أن تتغير إذا غيرنا من ترتيب السلع في المعادلة الجبرية سواء في البسيط أو في المقام . وفي الواقع نجد أن كل الصيغ تستوفي هذا الشرط .

اختبار الانعكاس في الزمن :

يقصد بذلك إذا كانت الصيغة المستخدمة في حساب الرقم القياسي يمكن أن تعمل في الاتجاهين تبعاً للزمن ؛ بمعنى انه إذا ارتفع سعر السكر من ليرة في سنة ١٩٦٠ وأصبح ٢ ليرة في سنة ١٩٦٥ – ارتفاع بنسبة ٢٠٠ ٪ – فان سعر السكر في سنة ١٩٦٠ إلى سنة ١٩٦٥ هو ٥٠ ٪ ، فكل من الرقمين هو مقلوب الآخر وبذلك يكون حاصل ضربهما يساوي واحد صحيح . وبالمثل إذا كانت نتيجة الرقم القياسي على أساس صيغة معينة = ٢٠٠ على اساس سنة ما ، فان هذه الصيغة لا بد أن تعطينا مقلوب هذا الرقم إذا عكسناها بمعنى ان الرقم القياسي في سنة الاساس منسوباًالىسنة المقارنة لا بد أن يكون ٥٠ . وبذلك لكي تكون الصيغة صحيحة يجب إذا ضربت في مقلوبها تعطينا ناتجاً = ١ صحيح ، أما اذا لم نحصل على هذه النتيجة فلا بد أن يكون هناك تحيز فيها والبديل الزمني لأية صيغة من صيغ الرقم القياسي هو نفس الرقم محسوباً بطريقة عكسية بحيث تعتبر سنة الاساس سنة مقارنة، وسنة المقارنة سنة أساس اي اننا نضع $ع_1$ بدل $ع_0$ وكذلك $ك_1$ بدل $ك_0$ اذا وجدت .

$$\text{فالرقم :}\quad \frac{\text{مج}\ ع_1}{\text{مج}\ ع_0}\quad \text{بديلة الزمن}\quad \frac{\text{مج}\ ع_0}{\text{مج}\ ع_1}$$

ونلاحظ ان ثلاث صيغ فقط تجتاز هذا الاختبار هي :

١ – الرقم التجميعي البسيط .

٢ – الرقم التحميعي المرجح بالوسط الحسابي لكميات سنتي الاساس والمقارنة

٣ – الرقم الأمثل .

(يمكن للطالب أن يطبق الانعكاس في الزمن كي يتأكد من ذلك)

اختبار الانعكاس في المعامل :

قيمة اية سلعة في فترة ما تساوي حاصل ضرب الكمية المنتجة من هذه السلعة في سعرها وهي تساوي جبرياً ع ك . وبذلك فالسعر والكمية هما معاملا القيمة ، والرقم القياسي للقيمة في سنة ما بالنسبة الى سنة اخرى

يساوي جبرياً $\frac{ع_1 ك_1}{ع_. ك_.}$

على هذا الاساس إذا كوّنا رقماً قياسياً لسعر سلعة ما (منسوب السعر) وكوّنّا كذلك رقماً قياسياً (منسوب) لكمية هذه السلعة في نفس المدة ، فمن المنطق أن يكون حاصل ضرب هذين الرقمين يساوي النسبة بين قيمتي هذه السلعة وإلا كان الرقم القياسي المستعمل بالنسبة للسعر والكمية خاطئاً في تصويره للتغير الذي يحدث في هاتين الظاهرتين .

وقياساً على ذلك إذا استعملنا أي صيغة من الصيغ التي نعرفها للأرقام القياسية في حساب الرقم القياسي للاسعار والرقم القياسي للكميات يجب أن يكون حاصل ضرب هذين الرقمين مساوياً للرقم القياسي لقيم هذه السلع في نفس المدة وإلا كانت الصيغة المستعملة خاطئة في تصويرها للتغير الذي يحدث في هاتين الظاهرتين .

هذا الاختبار يسمى الانعكاس في المعامل أو الانعكاس المعاملي ويتلخص

في أن أي صيغة للرقم القياسي إذا ضربت في بديلها المعاملي كانت النتيجة رقماً قياسياً للقيمة ، والبديل المعاملي لأية صيغة هو نفس الصيغة موضوعاً فيها ع بدل ك ، و ك بدل ع أي الكمية بدل السعر والسعر بدل الكمية . بمعنى انه إذا كان لدينا رقماً قياسياً للأسعار مرجحاً بالكميات فان البديل المعاملي يكون رقماً قياسيا للكميات مرجحاً بالأسعار . ويمكن وضع هـذا الشرط في الصيغة الجبرية الآتية :

$$\text{ان الرقم القياسي × البديل المعاملي له} = \frac{\text{مجـ ع}_{١}\text{ك}_{١}}{\text{مجـ ع}_{٠}\text{ك}_{٠}}$$

وبتطبيق هذا الاختبار يتضح لنا أن الرقم الأمثل هو وحده الذي يجتاز وهو لهذا السبب أحسن الارقام القياسية وأصحها من الناحية النظرية ومن ثم كانت تسميته بالرقم الأمثل .

ويمكن جبرياً اثبات انطباق هذا الشرط على الرقم الأمثل كالآتي :

$$\text{الرقم الأمثل} = \sqrt{\frac{\text{مجـ ع}_{١}\text{ك}_{٠}}{\text{مجـ ع}_{٠}\text{ك}_{٠}} \times \frac{\text{مجـ ع}_{١}\text{ك}_{١}}{\text{مجـ ع}_{٠}\text{ك}_{١}}}$$

$$\text{البديل المعاملي} = \sqrt{\frac{\text{مجـ ك}_{١}\text{ع}_{٠}}{\text{مجـ ك}_{٠}\text{ع}_{٠}} \times \frac{\text{مجـ ك}_{١}\text{ع}_{١}}{\text{مجـ ك}_{٠}\text{ع}_{١}}}$$

بضرب الرقم الأمثل في البديل المعاملي له نحصل على :

$$= \sqrt{\frac{\text{مح ع}_1\text{ ك}_0}{\text{مح ع}_0\text{ ك}_0} \times \frac{\text{مح ع}_1\text{ ك}_1}{\text{مح ع}_0\text{ ك}_1} \times \frac{\text{مح ك}_1\text{ ع}_0}{\text{مح ك}_0\text{ ع}_0} \times \frac{\text{مح ك}_1\text{ ع}_1}{\text{مح ك}_1\text{ ع}_1}}$$

$$= \sqrt{\frac{(\text{مح ك}_1\text{ ع}_1)^2}{(\text{مح ك}_0\text{ ع}_0)^2}}$$

$$= \frac{\text{مح ك}_1\text{ ع}_1}{\text{مح ك}_0\text{ ع}_0}$$ وهذا هو الرقم القياسي للقيمة .

تعديل الأرقام القياسية :

اتضح لنا أن هـــذه الاختبارات كانت قاسية بحيث لم تنطبق كل الشروط التي اقترحها فيشر الا على الرقم الأمثل فقط . ولما كان من الصعب أحياناً استخدام الرقم الأمثل نظراً لاحتمال عـــدم توفر البيانات اللازمة لتركيبه ، لهذا نحاول تعديل الأرقام القياسية التي لا تنعكس في الزمن أو في المعامل بحيث تصبح قابـلة للانعكاس ، ولكي نتوصل إلى ذلك نبحث عن معكوس الرقم .

والمعكوس الزمني للرقم هو خارج قسمـــة الواحد الصحيح على بديله الزمني فمثلاً الرقم التجميعي المرجح بكميات الأساس لا ينعكس في الزمن ومعادلته هي :

$$\frac{\text{مح ع}_1\text{ ك}_0}{\text{مح ع}_0\text{ ك}_0}$$ وبديله الزمني $$\frac{\text{مح ع}_0\text{ ك}_1}{\text{مح ع}_1\text{ ك}_1}$$

وبذلك يكون المقلوب الزمني $$\frac{\text{مح ع}_1\text{ ك}_1}{\text{مح ع}_0\text{ ك}_1}$$

وبذلك يكون الوسط الهندسي للرقم الأصلي ومقلوبه الزمني يقبل الانعكاس في الزمن. هذا الوسط الهندسي هو الرقم القياسي المعدل. وهكذا يمكننا تعديل أي رقم قياسي لكي ينعكس في الزمن بأن نوجد الوسط الهندسي بينه وبين مقلوبه الزمني.

كذلك يمكن تعديل الرقم القياسي حتى ينعكس في المعامل وذلك بايجاد الوسط الهندسي بين الرقم ومقلوبه المعاملي. والمقلوب المعاملي لأي رقم هو خارج قسمة منسوب القيم $\frac{\text{مجـ } \text{ع}_{١} \text{ك}_{١}}{\text{مجـ } \text{ع}_{٠} \text{ك}_{٠}}$ على البديل المعاملي للرقم نفسه.

فمثلاً الرقم المرجح بكميات سنة الأساس $\frac{\text{مجـ } \text{ع}_{١} \text{ك}_{٠}}{\text{مجـ } \text{ع}_{٠} \text{ك}_{٠}}$

بديله المعاملي $\frac{\text{مجـ } \text{ك}_{١} \text{ع}_{٠}}{\text{مجـ } \text{ك}_{٠} \text{ع}_{٠}}$

المقلوب المعاملي $= \frac{\text{مجـ } \text{ع}_{١} \text{ك}_{١}}{\text{مجـ } \text{ع}_{٠} \text{ك}_{٠}} \div \frac{\text{مجـ } \text{ك}_{١} \text{ع}_{٠}}{\text{مجـ } \text{ك}_{٠} \text{ع}_{٠}} = \frac{\text{مجـ } \text{ع}_{١} \text{ك}_{١}}{\text{مجـ } \text{ع}_{٠} \text{ك}_{١}}$

ويكون الوسط الهندسي للرقم الأصلي ومقلوبه المعاملي $= \sqrt{\frac{\text{مجـ } \text{ع}_{١} \text{ك}_{٠}}{\text{مجـ } \text{ع}_{٠} \text{ك}_{٠}} \times \frac{\text{مجـ } \text{ع}_{١} \text{ك}_{١}}{\text{مجـ } \text{ع}_{٠} \text{ك}_{١}}}$

وهذا هو الرقم الأمثل الذي نعرف انه ينعكس في المعامل.

وإذا كان الرقم لا ينعكس في الزمن أو في المعامل فيجب معالجته مرتين الأولى بالنسبة إلى الزمن والثانية بالنسبة إلى المعامل وسواء بدأنا بهذا أو بذاك تكون النتيجة رقماً يقبل الانعكاس في الزمن وفي المعامل.

ولكن هذا التعديل – مع الأسف – يعطينا في جميع الأحوال صيغاً

أكثر تعقيداً في التركيب وأشد ارهاقاً في الحساب من الصيغ الأصلية المقصود تعديلها وهذا التعقيد يقلل من فائدة هـــذا التعديل وأهميته من الناحية العملية .

الرقم القياسي للكميات :

يمكن حساب رقماً قياسياً للكميات بنفس الطرق السابقة مع ملاحظة **أن** تؤخذ الأسعار أو القيم كأوزان :

مثلاً $\frac{\text{مجـ ك}_{1}\ \text{ع}_{0}}{\text{مجـ ك}_{0}\ \text{ع}_{0}} \times ١٠٠$ رقم قياسي للكمية مرجحاً بأسعار الأساس .

أو $\frac{\text{مجـ ك}_{1}\ \text{ع}_{1}}{\text{مجـ ك}_{0}\ \text{ع}_{1}} \times ١٠٠$ رقم قياسي للكمية مرجحاً بأسعار المقارنة .

وهكذا بالنسبة لباقي الصيغ التي أشرنا اليها بالنسبة للأرقــام القياسية للاسعار .

كذلك يمكن حساب رقم قياس للكميات بطريقة المناسيب حيث يحسب منسوب الكمية لكل متغــير على حده ($\frac{\text{ك}_{1}}{\text{ك}_{0}} \times ١٠٠$) ثم بايجاد الوسط الحسابي او الوسط الهندسي او الوسط التوافقي (بسيطا او مرجحاً) ، لمناسيب المتغيرات التي نقرر ادخالها في الرقم القياسي نحصل على الرقم الذي نريده وهو مثل الرقم القياسي للاسعار يقيس التغير النسبي في كميات المتغيرات بالنسبة لفترة الاساس ويفضل الوسط الهندسي للمناسيب خاصة في الرقم القياسي لاسعار الجملة حيث انه يشمل عدداً كبيراً من السلع الــتي تختلف

اسعارها كثيراً ، وميزة الوسط الهندسي انه يقلل من تأثير القيم الكبيرة جداً والقيم الصغيرة جداً ، وذلك بعكس الوسط الحسابي الذي يتأثر بها كثيراً ، وفيما عـدا الرقم القياسي لأسعار الجملة لا يستخدم الوسط الهندسي نظراً للعمليات الحسابية المعقدة التي يحتاجها خـاصة إذا كان الرقم الذي نحسبه مرجحاً . وبذلك يكون الوسط الحسابي للمناسيب هو الوسط الشائع استعماله في حساب الارقام المختلفة .

استبعاد اثر تغيرات الاسعار :

بالاضافة إلى فائدة الرقم القياسي للأسعار في دراسة التغيرات النسبية في الاسعار على مر الزمن ، فانه يمكن بواسطة هذه الارقام معرفة التغيرات الحقيقية التي تحدث في الظواهر التي تتعرض قيمها لتضخم شديد بسبب ارتفاع الاسعار مثل الدخل القومي والاجور وقيم الصادرات والواردات ... الخ . ان التضخم الشديد في قيم هذه الظواهر يخفى تغيراتها الحقيقية التي قد نكون مهتمين بالبحث عنها . وبواسطة الارقام القياسية للاسعار نستطيع استبعاد اثر تغيرات الاسعار وبذلك لا يبقى في الظاهرة من تغير الا التغيرات الحقيقية أي تغيرات الكميات الخـاصة بمفردات الظاهرة والتي من الواضح لا يمكن جمعها سويا بسبب اختلاف وحداتها ويكون هذا الاستبعاد تطبيقاً للقاعدة ، الكمية × السعر = القيمة ، وبذلك فان قسمة القيمة على الرقم القيـاسي للاسعار يعطينا رقماً يمثـل التغير الحقيقي ، أي التغير في كمية الظاهرة أو الظواهر موضوع البحث . فاذا قسمنا قيمة الدخل القومي علي الرقم القياسي للاسعار نحصل على رقم معين يبين قيمة الدخل القومي مقاسية باسعار سنة الاساس . كذلك اذا قسمنا الرقم القياسي لكسب العمل على الرقم القياسي لنفقة المعيشة نحصل على رقم يمثل التغير النسبي في مستوى المعيشة (من الناحية المادية فقط) والذي نطلق عليه احياناً الرقم القياسي للاجر الحقيقي .

الارقـام القياسية الشائعة

الرقم القياسي لنفقة المعيشة :

يعتقد الاحصائيون في الوقت الحاضر ان تسمية هذا الرقم هكذا تسمية مضللة . ويفضلون تسميته بالرقم القياسي لاسعار التجزئه . ويقيس هــذا الرقم تكلفة شراء مجموعــة معينة من السلع التي يُعتقد انها تمثل استهلاك الطبقة التي من أجلها حسب هذا الرقم . ويحسب هذا الرقم بصيغة مرجحة باوزان سنة الاساس . وتعتمد الأوزان التي تستخدم في الترجيح على نتائج بحث ميزانية الاسرة . والطريقة الشائعة في حساب هذا الرقم هي على اساس الوسط الحسابي المرجح لمناسيب الاسعار .

وفي الغالب يحسب هذا الرقم شهرياً ، حيث تجمع الاسعار في يوم معين في منتصف الشهر. وفي بعض الدول تجمع الاسعار اسبوعياً ثم يحسب المتوسط لهـا ، وتكون نتيجة الرقم بذلك اكثر دقة ، إلا ان ذلك يتطلب جهداً أكبر ، خاصة اذا كان الرقم يتضمن عددا كبيراً من السلع .

وتقوم مصلحة الاحصاء في أغلب الدول بجمع الاسعار اما بارسال موظفيها إلى محلات معينة او بمطالبة اصحاب هذه المحلات بالتبليغ عن اسعار السلع المحـــددة مرة كل شهر . والمشكلة الاساسية في هذه العملية هي ضرورة التأكد من ان الاسعار التي تجمعها المصلحة تمثل اسعار نفس السلع كل شهر .

الرقم القياسي لاسعار الجملة :

يقيس هذا الرقم التغيرات النسبية في اسعار السلع التي يتم التبادل فيها في اسواق الجملة . وكان الاحصائيون يأملون في بادىء الأمر الحصول عـلى مقياس يمكن ان يقيس التغيرات العامة في الاسعار ، إلا انه في الوقت الحاضر أصبح التركيز على مقياس لتغيرات اسعار مجموعات معينة من السلع .

ويحسب هذا الرقم بصيغة مرجحة ، حيث تستخدم قيم السلع التي جرى تبادلها في فترة الأساس . إلا انه في بعض الدول يجري ترجيح أسعار السلع بأخذ أسعار عدد معين من أصناف كل سلعة يتناسب مع أهميتها ، بمعنى انه كلما أردنا إعطاء السلع وزناً إكبر ندخل في الرقم عدداً أكبر من أصنافها .

وفي الغالب يحسب الرقم القياسي لأسعار الجملة بطريقة الوسط الهندسي المرجح للمناسيب خاصة إذا كنا نتبع نظاماً بسيطاً للترجيح (مثلا أخذ أسعار عدد معين من أصناف كل سلعة) .

الرقم القياسي للانتاج الصناعي :

هذا رقم قياسي الكميات يهدف إلى قياس التغيرات في انتاج مجموعة معينة من الصناعات . وفترة الأساس لهذا الرقم هي غالباً سنة ، والأوزان هي القيمة الصافية لانتاج كل سلعة في فترة الأساس ويحسب الرقم بطريقة الوسط الحسابي المرجح لمناسيب الكميات .

وتجمع المعلومات الخاصة بهذا الرقم بمطالبة عينة من المؤسسات بالتبليغ شهريا عن إنتاجها . وفي بعض الحالات التي لا يمكن فيها جمع بيانات عن الإنتاج تؤخذ معلومات اخرى تكون مؤشر على الانتاج وذلك مثل البيانات الخاصة بالتوظف ، على ان ذلك لا يجب أن يتبع الا اذا يئسنا من الحصول على بيانات خاصة بالإنتاج نفسه حيث ان بيانات التوظف تجعل الرقم لا يصلح كمقياس الإنتاجية .

الارقام القياسية لحجم الصادرات والواردات .

الهدف من هذه الارقام هو قياس التغيرات في قيم الواردات والصادرات بأسعار ثابتة . ان التغير في قيم الواردات او الصادرات يمكن أن يكون نتيجة تغير كميات السلع المستوردة أو المصدرة او نتيجة تغيرات اسعارها او

نتيجتهما سويا . والهدف من هذه الارقام القياسية هو استبعـــاد اثر تغيرات الاسعار حتى يمكن ملاحظة التغيرات الحقيقية في حجم التجارة .

ويحسب هذا الرقم بطريقة الوسط الحسابي المرجح لمناسيب الكميات ، حيث تستخدم قيم السلع في فـــترة الاساس كأوزان ، وفـــترة الأساس هي غالباً سنة وتحسب المناسيب على أساس المعلومات التي تظهرها النشرات الاحصائية للتجارة الخارجية حيث نجد فيها بيانات عن قيم وكميات السلع المستوردة او المصدرة كل شهر . وفي بعض الحالات عندمـــا لا نتوفر بيانات دقيقة ومناسبة عن الكميات يمكن التوصل الى رقم قياسي للكميات بقسمة رقم القيم على رقم الأسعار الخاصة بالبضائع التي جرى التبـــادل فيها والذي يحسب خصيصاً لهذا الغرض وفي هذه الحالة يكون رقم الكميات الناتج رقماً مرجحاً بقيم فترة المقارنة وليس قيم فترة الاساس . ومن الواضح أنه في هذه الحالة يحسن استخدام صيغة فيشر لحساب الرقم القياسي للأسعار.

وإذا حسبنا الرقم القياسي لأسعار الصادرات أو الواردات فإنـــه يظهر نسب التغيرات في قيم التجارة التي ترتبت على تغيرات الأسعار ، فإذا أظهر ارقم القياسي لأسعـــار الواردات ارتفاعاً بنسبه ٢٠ ٪ مثلاً فإن ٢٠ ٪ من الزيادة في قيم الواردات تكون راجعة إلى هذا التغير في أسعارها . وإذا نوفر لدينا رقما قياسياً لقيم الواردات أو الصادرات ورقما قياسياً لكمياتها فبقسمة الأول على الثاني نحصل على رقم قياسي لأسمارها .

ومن الواضح أن مقـــارنة الرقم القياسي لأسعـــار الصادرات بالرقم القياسي لأسعار الواردات تدلنا على معدل التبادل في التجارة الخارجية وهو مقياس ذو أهمية كبرى في تحديد الكسب الذي يعود على الدولة من تجارتها الخارجية .

تمارين

١) الآتي بيانات عن انتاج احدى الدول من القمح بالطن :

السنة	١٩٥٠	١٩٥١	١٩٥٢	١٩٥٣	١٩٥٤	١٩٥٥	١٩٥٦
الانتاج	١٠١٩	٩٨٨	١٣٠٦	١١٧٣	٩٩٤	٩٣٥	١٠٠٤

احسب مناسيب الانتاج على أساس عام ١٩٥٥

» » » » » متوسط الأعوام من ١٩٥٠ – ١٩٥٣ .

٢) إذا علمت أن منسوب الكمية لعام ١٩٥٨ على أساس عام ١٩٤٩ = ١٠٥ وان منسوب الكمية لعام ١٩٥٨ على أساس ١٩٥٣ = ١٤٠ – احسب منسوب الكمية عام ١٩٥٣ على أساس عام ١٩٤٩ .

٣) تتوقع احدى المؤسسات زيادة مبيعاتها من سلعة مـا بنسبة ٥٠٪ في العام القادم . بأي نسبة يجب أن ترفع سعر هذه السلعة حتى يمكن أن تزيد ايراداتها إلى الضعف .

٤) إذا كانت مناسيب الأسعار في الفترة من عام ١٩٥٦ إلى عام ١٩٦٠ بطريقة السلسلة هي على التوالي ١٢٥ ، ١٢٠ ، ١٣٥ ، ١٥٠ ، ١٧٥ . حول هذه المناسيب إلى مناسيب على أساس عام ١٩٥٦ .

٥) من البيانات الآتية احسب الرقم القياسي للأسعار بصيغـــة فيشر وكذلك الرقم القياسي للكميات (١٩٥٠ على أساس ١٩٤٩) .

السلع	اسعار ١٩٤٩	اسعار ١٩٥٠	كميات انتاج ١٩٤٩	كميات انتاج ١٩٥٠
ا	٣ر٩٥	٣ر٨٩	٩٦٧٥	٩٧١٧
ب	٦١ر٥	٦٢ر٢	١١٧٧	١١٥٥
جـ	٣٤ر٥	٣٥ر٤	٧٧٩٣	٧٥٣٩

٦) إذا علمت أن متوسط اجر العامل في السنوات من عام ١٩٤٧ إلى عام ١٩٥١ هو ١٩ر١ ، ٣٢ر١ ، ٤٤ر١ ، ٥٧ر١ دينار . وان الأرقام القياسية للاسعار هي على التوالي ١٠٠ ، ٦ر١٠٧ ، ٦ر١٠٦ ، ٦ر١٠٧ . احسب متوسط اجر العامل باسعار عام ١٩٤٧ .

الفصل الرابع عشر

السلاسل الزمنية

مقدمة :

يهتم الاقتصاديون بنوعين من الدراسات الاقتصادية ، النوع الأول هو الذي يتجه إلى تحليل حالة النشاط الاقتصادي التي يمكن ان توجد إذا كانت القوى الاقتصادية في توازن تام بحيث لا يصيب النشاط الاقتصادي أي اختلال ، هذا النوع من التحليل يطلق عليه الاقتصاديون التحليل الاستاتيكي . والنوع الثاني وهو الذي يطلق عليه التحليل الديناميكي يتجه إلى دراسة ما يمكن أن يصيب النشاط الاقتصادي من اختلال نتيجة تغير القوى الاقتصادية المختلفة فبينما نهتم في النوع الأول بدراسة الحالة التي يصل اليها النشاط الاقتصادي بعد الوصول إلى التوازن نهتم في النوع الثاني بتحليل ما يمكن أن يحدث من تغييرات قبل الوصول إلى التوازن .

والواقع أن الاقتصاديين في العصر الحديث أصبحوا يركزون اهتمامهم على النوع الثاني من التحليل الاقتصادي حيث انهم تنبهوا إلى أن حالة التوازن ليست إلا حالة نظرية لا يمكن أن يقف عندها النشاط الاقتصادي الذي يبدو أنه دائم التغيير والحركة . هذا النوع من الدراسة يقوم إلى حد بعيد على الدراسة الاحصائية للسلاسل الزمنية ، تلك الدراسة التي كان للاقتصاديين الفضل الكبير في تقدمها تقدماً كان من نتيجته ابتداع الأساليب الاحصائية

التي يمكن استخدامها في تحليل ما يصيب النشاط الاقتصادي من تغيير مع الزمن ، وهي أساليب تختلف تماماً عن تلك التي تستخدم في تحليل التوزيعات التكرارية التي تدرس الظواهر في حالة السكون . ولو ان الاقتصاديين كان لهم الفضل الأكبر في الدراسة الاحصائية للسلاسل الزمنيــة الا ان فائدتها ليست قاصرة عليهم فهي تهم كذلك رجال الأعمال ، والباحثين الاجتماعيين والأطباء وعلماء الحياة والعاملين في ميادين الصحة العامة .

تعريف السلسلة الزمنية :

يمكننا أن نعرف السلسلة الزمنية لاية ظاهرة بانها التطور التاريخي لهذه الظاهرة ، وبمعنى آخر هي القيم أو المقادير التي تتخذهـــا هذه الظاهرة في فترات زمنية متتابعة قد تكون أيام أو أسابيع أو أشهر أو سنين . ولتركيب سلسلة زمنية لظاهر ما نعمل أولاً على مشاهدتها مدة من الزمن ونسجل أثناء المشاهدة القيم المختلفة التي تتخذها في فترات زمنيـــة منتظمة فيمكننا مثلاً أن نكوّن سلسلة زمنية لكمية البنكنوت المتداول في دولة ما بسؤال البنك المركزي في هذه الدولة أو بالرجوع إلى الاحصاءات التي ينشرها عن هذا الموضوع فنحصل على قيم شهريه نضعها في جدول يبين بـــه الفترات الزمنية والقيم المقابلة للبنكنوت المتداول ، هذا الجدول يوضح السلسلة الزمنية لهذه الظاهرة خلال هذه المدة من الزمن .

والهدف من الدراسة الاحصائية للسلاسل الزمنية هو الكشف عن التغيرات التي تطرأ على الظواهر التي ندرسها أثناء مدة معينة حتى يمكن معرفة أنواع هذه التغيرات وقياس كل نوع منها . فالتحليل الاحصائي لأية سلسلة زمنيـــة يقوم أساساً على مقارنة قيم الظاهرة في فترات متتابعة حتى يمكن الكشف عمّا يصيبها من نمو أو ضمور . ولذلك يجب أن نتنبه الى ان المقارنة لكي تكون صحيحة يجب أن تقاس الظاهرة بنفس الوحدات وبنفس الطريقة في الفترات

الزمنية المتتابعة بمعنى أن لا تتغير طبيعتها أو صفاتها التي تميزها أو على الأقل لا تتغير بدرجة كبيرة .

الدراسة الاحصائية للسلسلة الزمنية :

تبدأ الدراسة الاحصائية لأي سلسلة زمنية برسم خط بيـــاني لها يوضح تغيرها مع الزمن ، فنبين الزمن على المحور الأفقي والقيم أو المقادير على المحور الرأسي ، ثم نضع نقط أمام كل فترة زمنية والقيمة المقابـــلة لها ، ثم نصل هذه النقط فيتكون لدينا خط بياني يسمى أحيانا بالمنحنى التاريخي للظاهرة التي ندرسها .

بعد ذلك تتجه الدراسة الإحصائية للسلاسل الزمنية الى تحليل التغييرات التي تطرأ على الظواهر موضوع البحث إلى أنواعها المختلفة ثم قياسها وتحديد اتجاهها سواء كان نحو الزيادة أو نحو النقص والاستفادة من ذلـــك في فهم طبيعة هذا التغيرات فهما لـكي نتوصل اليه يجب ان يكون لدينا معرفة تامة بموضوع البحث .

كما ان دراسة السلاسل الزمنية تهدف أيضاً الى الاستفـــادة من دراستنا لماضي الظواهر في عمل تقديرات لها في المستقبل حتى يمكن أن نستعد لمواجهة ما يطرأ عليها من تغييرات فلا نفاجأ بها . كذلك يتضمن التحليل الاحصائي للسلاسل الزمنية دراسة ما يمكن ان يوجد من ارتباط بين سلسلتين أو أكثر.

أنواع التغيرات التي يمكن أن تحتويهـــا السلسلة الزمنية :

ليس هناك اتفاق تام بين الاقتصاديين على الأنواع المختلفة للتغيرات التي يمكن أن تظهر في السلسلة الزمنية . إلا أن هناك اربع انواع رئيسية يـــكاد الجميع يجمع على تمييزها ودراستها وهي التغيرات الموسمية والدورية والعرضية والاتجـــاه العام .

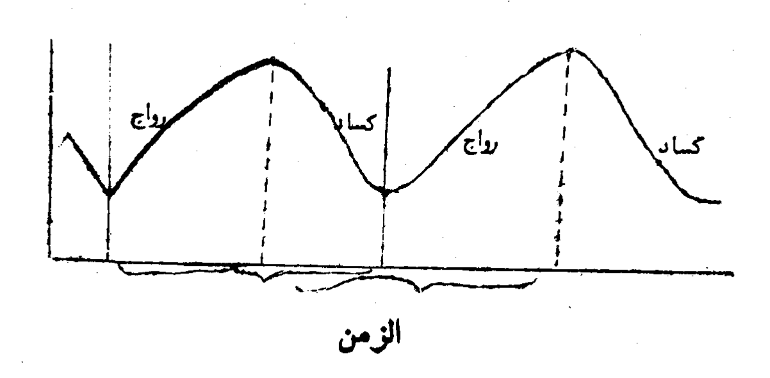
١ – **التغيرات العرضية** – وهي تغيرات لم يحاول الاقتصاديون وضع نظرية لتفسيرها حيث انها تنشأ عن اسباب عارضة لم تكن في الحسبان ولذا فهي شاذة وطارئة من الوجهة النظرية بمعنى انه لا يمكن التنبؤ بوقوعها أو تحديد مقدارها حيث انها لا تتبع أي قاعدة او قانون . ويترتب على ذلك اننا لا ننتظر تكرار هذا النوع من التغير وانما هو يفاجئنا . ومثل هذه التغييرات ما يصيب النشاط الاقتصادي نتيجة زلزال أو فيضان او حرب أو اضراب عام بين العمال .

على أن هذه التغيرات قد تكون من الضآلة بحيث لا تستحق الانتباه اليها وقد تكون من الضخامة بحيث تثير انتباهنا ، إذ قد يترتب على هذه الضخامة نشوء نوع آخر من التغير ، فالحرب مثلا تعتبر في بدايتها تغيراً عرضياً ولكنها بعد ذلك تصبح تغيراً دورياً له أهميته .

٢ – **التغيرات الموسمية** – وهي تغيرات منتظمة انتظاماً تاما إذا قورنت بأي نوع آخر من التغيرات ، فهي تغيرات تكرر في فترات منتظمة بحيث تحدث في تواريخ معلومة في كل سنة ولا تحيد عنها بحيث يكون اتجاهها واحد لا يتغير . فكميه النقد المتداول في مصر تتجه نحو الزيادة خلال شهر سبتمبر واكتوبر ونوفمبر من كل عام وإلى النقص خلال شهر يوليو وأغسطس ، والمهم في هذه التغيرات ان أسبابها ليست كامنة في تفاعل النشاط الاقتصادي نفسه ، وانما هي أسباب مختلفة عن ذلك قد تكون التقاليد الاجتماعية مثل الأعياد وقد يكون تغير الجو مثل موسم الصيف في لبنان وقد يكون ظهور المحصول الهام في الدولة في فترة زمنية معينة مثل القطن في مصر .

وحيث أن هذهالتغيرات تتصف بالتكرار فلذا يمكن دراستها احصائياً حتى يمكن قياسها ومعرفة اتجاهها والاستعداد لهـا في المستقبل . والدراسة الاحصائية لهذه التغيرات تقوم على أساس مشاهدة الظاهرة التي ندرسها في الفترات المتتابعة التي تتكون منها السنة مثل الأشهر او ارباع السنة .

٣ - **التغيرات الدورية** - وهي تغيرات منتظمة ولكن انتظامها ليس تاماً كما هو الحل في التغيرات الموسميـة ، إذ أن طول الدورة الاقتصادية ليس معروفاً بدقـــة كما انه يتغير ولذلك لا يمكن تحديد مواعيد هذه التغيرات بدقة ، والتغيرات الدورية ليست تغيرات قصيرة الأجل مثل التغيرات الموسميه إذ لا يظهر لها أثر محسوس بين سنة وأخرى وانما تستغرق زمناً طويلاً حتى تستعيد سيرتها وقد اختلف الاقتصاديون في تحديـد هذا الزمن ولكن الغـالب أن الدورة تستغرق حوالي عشر سنوات ، والمقصود بطول الدورة المدة التي تمضي بين قمة الموجـه والقمة التي تليها (قمة الرواج) أو بين قاع الموجـه والقاع الذي يليه (قاع الكساد) كما يتضح من الرسم التالي .



وتختلف هذه التغيرات كذلك عن التغيرات الموسمية في أن أسبابها اقتصادية بحتة بمعنى انها كامنة في تفاعل النشاط الاقتصادي نفسه وقد وضعت نظريات اقتصادية كثيرة تحاول تفسير هذه الأسباب (نظرية كينز مثلاً) .

وبالطبع يمكن دراسة هذه التغيرات دراسة احصائية لأنها كما ذكرنا تتصف بالتكرار . والدراسة الاحصائية لها تقوم على أساس مشاهدة الظاهرة التي نبحثهـــا في مدة طويلة ، عشرين سنه مثلاً .

٤ – **الاتجاه العام** – أثناء حدوث الأنواع السابقة للتغيرات يطرأ على الظاهرة تغير بطيء لا يمكن ملاحظته في المــدة القصيرة وانما يكون واضحاً بعد مرور مدة طويلة من الزمن ، هذا التغير هو ما نسميه بالاتجاه العام وهو تغـــير طبيعي يطرأ على الظواهر المختلفة فيوضح ما يصيبها من نمو في فترة تاريخية معينة ومن ضمور في فترة تاريخية أخرى . فاذا درسنا تطور صناعة ما في دولة معينة يمكن أن نلاحظ تغيراً بطيئاً يصيبها تدريجياً بحيث تأخذ في النمو حتى تصل إلى قمة معينة ثم يطرأ عليها تغير تدريجي بطيء أيضاً في اتجاه عكسي بحيث تأخـــذ في الضمور بمرور الزمن . أما عن أسباب هذا النوع من التغير البطيء التدريجي بالنسبة للمسائل الاقتصادية فغير واضحة تماماً وهناك آراء كثيرة في هذا الصدد .

تفسير التغيرات التي تتضمنها السلسلة الزمنية :

لا يجب أن نعتقد أن الأنواع السابقـــة من التغيرات تحدث منفصلة عن

بعضها بحيث يمكن القول ان الظاهرة في هذه الفترة تخضع لتغير موسمي فقط وفي تلك لتغير عرضي فقط وهكذا ، ذلك لان هذه الانواع تحدث جميعاً في وقت واحد ، فيعمل كل نوع من ناحيته على التأثير على الظاهرة بدرجة معينة وفي اتجاه معين . وبذلك يكون أي تغير يطرأ على الظاهرة هو في الواقع المحصلة لجميع القوى والمؤثرات التي تحيط بالظاهرة ، فحالة الرواج مثلاً تقوي من التغير الموسمي الذي يكون في اتجاه صعودي ، وتضعف من التغير الذي يكون في اتجاه نزولي ، والكساد على العكس يقوي من التغير الموسمي النزولي ويضعف من التغير الصعودي . والاتجاه العام نحو الصعود يضعف من الكساد ويقوي من الرواج والعكس بالنسبة للاتجاه العام نحو الضمور فانه يقوي من الكساد ويضعف من الرواج .

ولذلك تنشأ لدينا عند الدراسة الاحصائية للسلاسل الزمنية مشكلة **تحليل** التغير الذي يطرأ على الظاهرة إلى أنواعه الأربعة وبمعنى آخر مشكلة **الفصل** بينها حتى يمكن تحديد اتجاه كل منها وقياسه .

قياس التغيرات وتحديد اتجاهها :

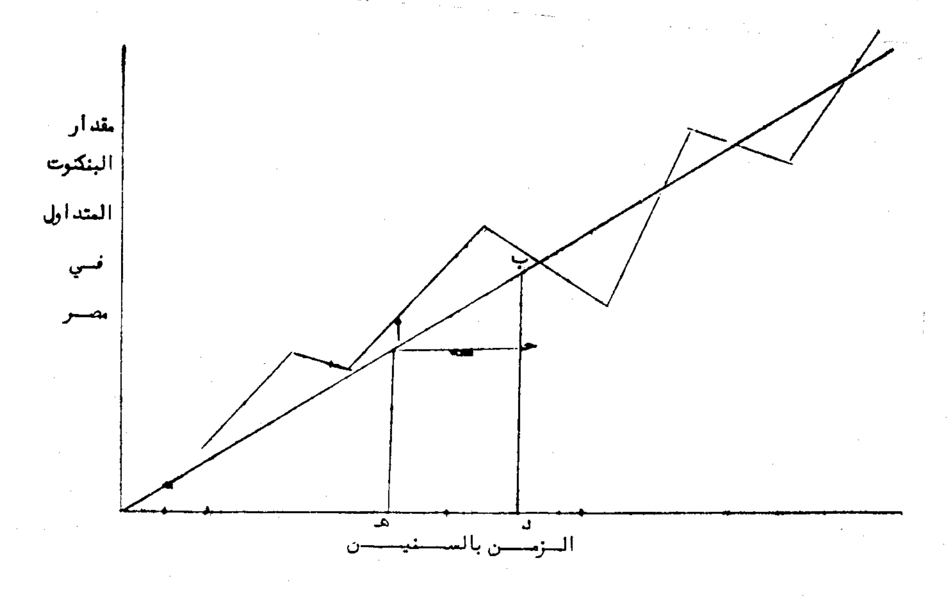
سبق ان ذكرنا أن التغيرات العرضية تحدث دون قاعدة أو قانون ولذا لا يمكن قياسها أو تقديرها حتى تحصل فعلاً . ولكن لكي نستطيع أن نقيس الانواع الأخرى من التغيرات يجب أن نعمل على استبعاد التغيرات العرضية ويتسنى لنا ذلك بمشاهدة الظاهرة خلال مدة طويلة من الزمن حتى يمكن لهذا النوع من التغيرات أن يمحو بعضه بعضاً فتنكشف لنا الاتجاهات الحقيقية للظاهرة .

ولقياس الاتجاه العام نحسب مقدار التغير الذي يصيب الظاهرة أثناء وحدة الزمن في المتوسط أي بعد استبعاد تأثير التغيرات العرضية ، ذلك لاننا نقصد بالاتجاه العام معدل التغير في المتوسط بالنسبة لوحدة الزمن ،

ويجب ذكر عبارة « في المتوسط » حيث نقصد بها اننا استبعدنا أثر التغيرات العرضية والموسمية والدورية .

ويستبعد أثر التغيرات الموسمية بأن نأخذ القيم السنوية للظاهرة - مجموعها أو متوسطها - لأن هذه التغيرات تستعيد سيرتها كل ١٢ شهراً . أما أثر التغيرات العرضية فيستبعد بتمهيد الخط البياني الذي يمثـل السلسلة الزمنية للظاهرة أو بطريقة المتوسطات المتحركة التي سيأتي الكلام عنها فيما بعد .
وتمهيد الخط البياني للسلسلة يجعلنا نحصل على خط يسمى بخط الاتجـاه العام وهو إما أن يكون مستقيما أو غير مستقيم . فإذا كان مستقيما دل ذلك على ان معدل التغير يسير بانتظام خلال الفترة التي ندرسها ، وإذا كان غير مستقيم دل ذلك على ان معدل التغير (الاتجاه العام) ليس منتظما خلال هذه الفترة بمعنى انه قد يكون صعوديا في بادىء الامر ثم يغير من اتجاهه فيصبح نزولياً بعد مرور مدة من الزمن .

بعد تمهيد الخط البياني للسلسلة نحدد الفترة التي نريد قياس الاتجاه العام فيها ثم نحدد نقطتين على خط الاتجاه العام بحيث يقعا خلال هذه الفترة ومنها نسقط عمودين على المحور الأفقي ومن النقطة الأولى نسقط عمـوداً موازيا للمحور الأفقي . ثم نحسب ظل الزاوية التي يصنعها خط الاتجاه العام مع هذا المحور فيكون هو معدل التغير الذي يميز تبعاً لوحدات الـتي قيست بها الظاهرة والذي يدل على مقدار التغير في المتوسط في وحدة الزمن .



واضح من الرسم كيف مهدنا خط الاتجاه العام بحيث يكون خالياً من التعريجات غير المنتظمة أو الذبذبات قصيرة الاجل بحيث يصور الاتجاه العام للظاهرة التي ندرسها تصويراً دقيقاً على هذا الخط حددنا نقطتين : أ ، ب ومنها أسقطنا عمودين أ هـ ، ب د . ومن النقطة أ أسقطنا عمود أ د على ب د . ومعدل التغير يكون بذلك هو ظل الزاوية ب أ حـ ويساوي $\frac{\text{ب حـ}}{\text{أ حـ}}$

وحيث ان خط الايجاه العام مستقيم فلا يتغير معدل التغير مهما غيرنا من النقط التي نحددها على هذا الخط .

وبنفس الطريقة يمكننا تمهيد خط غير مستقيم ، إلا أنه في هذه الحالة يكون هناك اتجاهان أحدهما صعوديا والآخر نزولياً فيكون ظل الزاوية في الحالة الأولى موجباً أي يكون معدل التغير بالزيادة بينما يكون ظل الزاويةفي الحالة الثانية سالباً أي يكون معدل التغير بالنقص . ونلاحظ في حالة الاتجاه غير المستقيم أن الفترة الزمنية التي نأخذها لحساب معدل التغير يجب

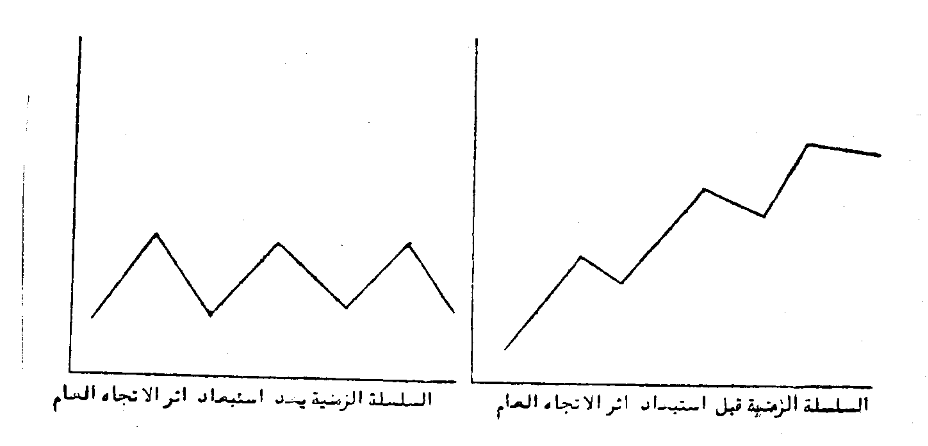
ان تكون فترة قصيرة جداً خوفاً من أن يتغير الاتجاه بمقدار كبير إذا أخذنا فترة زمنية طويلة .

على ان معدل التغير الذي نحسبه بالطريقة السابقة يكون عرضة للحكم الشخصي الذي يتوقف الى حـــد بعيد على رسم خط الاتجاه العام وليس من السهل أن يتفق أكثر من شخص واحد في رسم الاتجاه العـــام لسلسلة زمنية فيها ذبذبات كثيرة وعنيفة . ولهذا نلجأ إلى طريقة اخرى كي نضعف من التعريجات التي تتخلل المنحنى التاريخي بحيث يصبح هذا المنحنى قريباً من التمهيد فيكون من السهل بعد ذلك رسم خط الاتجاه العام . وتقـــوم هذه الطريقة على أساس المتوسطات المتحركة وذلك بأن نأخذ متوسطات القيم خلال عدة فترات ونعتبر هذا المتوسط كأنه قيمة الظاهرة في منتصف هذه المدة ثم نترك الفترة الاولى ونأخذ فترة اخرى من أسفل ونحسب المتوسط وهكذا ، فيصبح لدينا قيم ليس فيها أثر للتغيرات العرضية القصيرة الأجل أي خالية من الذبذبات العنيفة التي قد تكون في القيم الأصلية وبذلك يكون من السهل تحديد خـــط الاتجاه العام لها ، الذي نحسب على أساسه معدل التغير أي الاتجاه العام للظاهرة بنفس الطريقة السابقة .

ونلاحظ أنه كلما طالت المدة التي نحسب على أساسها المتوسطات المتحركة كلما صارت السلسلة الزمنية اكثر تمهيداً . على ان ذلك يؤدي إلى قصر السلسلة الزمنية الجديدة التي تتكون من المتوسطات المتحركة وهذا عيب إذ ان ذلك قد يؤدي الى ضياع معالم السلسلة الزمنية الأصلية حيث يصبح اتجاهها العام مختلفاً بعض الشيء عن الاتجاه العام للمتوسطات المتحركة . كذلك نلاحظ انه إذا كانت السلسلة الزمنية تتضمن تغيرات موسمية فلكي تخلص السلسلة من هذه التغيرات يجب أن تحسب المتـــوسطات المتحركة على اساس طول الفترة التي يحدث فيها التغير الموسمي وهو ١٢ شهراً او الاجزاء التي تتكون منها السنة والتي تكون السلسلة الزمنية الاصليـــة مقاسة تبعاً لها ..

وهناك طريقة ثالثة يمكننا ان نلجأ اليها لتمهيد السلسلة الزمنيـة وذلك بتوفيق المنحني الذي يمثل هذه السلسلة اي بتوفيق المعادلة الرياضية التي تمثل العلامة بين قيم الظاهرة والزمن . والزمن هنا يعتبر المتغير المستقل وقيم الظاهرة هي المتغير التابع . والتمهيد بهذه الطريقة ادق من التوفيـــق بالطريقتين السابقتين ولكن العيب الذي يؤخذ على هذه الطريقة هو صعوبتها من الناحية العملية خاصة اذا كان عدد القيم كبيراً او اذا كانت المعادلة التي نوفقها تمثل العلاقة بين الزمن والظاهرة خلال المدة المشمولة بالبيانات التي لدينا فقط بمعنى انه لا يمكن تطبيق هذه المعادلة على الماضي او المستقبل الا اذا فرضنا ان الظروف التي احاطت او ستحيط بهذه الظاهرة لا تختلف كثيراً عن الظروف التي احاطت بها اثناء المدة التي درسناها .

بعد حساب الاتجاه العام اي معدل التغير يمكن بطرحه طرحاً جبرياً اي بالإحتفاظ باشارته الجبرية من القيم الاصلية للظاهرة ان نحصل على قيم جديدة خالية من تأثير هذا الاتجاه العام ولكنها تتضمن طبعاً الانواع الاخرى من التغيرات ويمكن تصوير هذه العملية بالرسم البياني الآتي :



السلسلة الزمنية قبل استبعاد اثر الاتجاه العام — السلسلة الزمنية بعد استبعاد اثر الاتجاه العام

وبذلك يمكنا ان نحدد الطرق الآتية لتقدير الاتجاه العـام لظاهرة ما ،

اي لتقدير متوسط التغير في الظاهرة في وحدة الزمن .

١ – تحديد خط الاتجاه العام برسمه على السلسلة ثم حساب ميل الخط من الرسم البياني . ومن الواضح ان لهذه الطريقة عيب هام حيث ان النتيجة تتوقف على قدرة الانسان على تحديد الاتجاه العام للسلسلة بدقة .

٢ – تمهيد الذبذبات التي تظهر في السلسلة بحساب المتوسطات المتحركة التي يمكن بواسطتها استبعاد اثر التغيرات وبرسم السلسلة من واقع هذه المتوسطات يكون من السهل تحديد خط الاتجاه العام لها ثم حساب ميله الذي يدل على قيمة الاتجاه العام . وهذه الطريقة وان كانت اكثر دقة من الطريقة السابقة الا انها لا زالت تعتمد على حرية الانسان في رسم خط الاتجاه العام .

٣ – تمهيد السلسلة بايجاد معادلة خط الاتجاه العام سواء كان مستقيما او منحنى وذلك باتباع طريقة المربعات الصغرى. وفي تطبيق هذه الطريقة يمكن اعتبار الفترة الاولى هي نقطة الاصل في الرسم وبذلك يعبر عنها بالرقم صفر ثم الفترة التالية بالرقم ١ وهكذا ، كما يمكن اخذ فترة متوسطة واعتبارها هي فترة الاساس اي نقطة الاصل ويعبر عنها بالرقم صفر وعن الفترات السابقة لها بالارقام – ١ ، – ٢ ، – ٣ وهكذا ابتداء من نقطة الاصل وعن التغيرات اللاحقة لها بالارقام + ١ ، + ٢ ، + ٣ وهكذا ابتداء من نقطة الاصل . واتباع الطريقة الثانية يكون افضل حيث أنه يسهل علينا العمليات الحسابية خاصة اذا كان عدد الفترات فرديا . ونلاحظ ان قيمة الاتجاه العام وهي ميل الخط الذي يمثله معادلة الاتجاه العام (القيمة م في المعادلة) ، لا تختلف سواء اتبعنا الطريقة الاولى او الثانية ، الا ان قيمة ب تتغير ولذلك بعد ايجاد المعادلة يجب ان نتبعها بملاحظة نحدد فيها السنة التي اعتبرت نقطة الاصل للمعادلة .

٤ – يمكن ان نقسم السلسلة الى نصفين ثم نحسب متوسط كل نصف وبذلك

نحصل على قيمتين نحددهما على الرسم الخاص بالسلسلةثم نصل بينهما بخط مستقيم نحسب ميله لايجاد قيمة الاتجاه العام . ومن الواضح ان هذه الطريقة لا يمكن اتباعها الا اذا كان الاتجاه العام مستقيما .

ونلاحظ أنه اذا أخذنا متوسطات متحركةعلى أساس فترة زوجية (اربع سنوات مثلا) فان القيمة المتوسطة لا تكون سنة معينة وهذا ما يجعل التمهيد قليل الفائدة حيث يكون هدفنا هو الحصول على القيم الاتجاهية (أي قيم الظاهرة وهي تتأثر فقط بالاتجاه العام أي قيمتها بعـد استبعادات التغيرات الاخرى) المقابلة لكل سنة لذلك يمكن أخذ متوسط كل متوسطين الأول والثاني ، ثم الثاني والثالث وهكـذا وبذلك نحصل على متوسطات تقابـل سنوات معينة

مثال ٦٩ :

السنوات	القيم	متوسط متحرك لكل خمس سنوات
١٩٤٨	٥٠ر٠	
١٩٤٩	٣٦ر٥	
١٩٥٠	٤٣ر٠	٤٢ر٦
١٩٥١	٤٤ر٥	٤٠ر٢
١٩٥٢	٣٨ر٩	٣٩ر٤
١٩٥٣	٣٨ر١	٣٩ر٦
١٩٥٤	٣٢ر٦	٣٨ر٠
١٩٥٥	٣٨ر٧	٢٨ر٤
١٩٥٦	٤١ر٧	٣٧ر٦
١٩٥٧	٤١ر١	
١٩٥٨	٣٨ر٨	

السنوات	القيم	متوسط متحرك لكل اربع سنوات	متوسط متحرك لكل متوسطين
١٩٤٨	٥٠ر٠		
١٩٤٩	٣٦ر٥		
١٩٥٠	٤٣ر٠	٤٣ر٥	٤٢ر١
١٩٥١	٤٤ر٥	٤٠ر٧	٤٠ر٩
١٩٥٢	٣٨ر٩	٤١ر١	٣٩ر٨
١٩٥٣	٣٨ر١	٣٨ر٥	٣٧ر٨
١٩٥٤	٣٢ر٦	٣٧ر١	٣٨ر٥
١٩٥٥	٣٨ر٧	٣٧ر٨	٣٨ر٢
١٩٥٦	٤١ر٧	٣٨ر٥	٣٨ر٧
١٩٥٧	٤١ر١	٣٨ر٨	
١٩٥٨	٣٣ر٨		

وبرسم السلسلة من واقع المتوسطات الأخيرة سواء على أساس خمس سنوات أو على أساس أربع سنوات نحصل على خط بياني ممهد تقريباً وبذلك يكون من السهل تحديد اتجاهه العام وحساب ميله .

مثال ٧٠ :

ايجاد الاتجاه العام بتوفيق المستقيم بطريقة المربعات الصغرى :

السنوات	القيم (ص)	س	س^٢	س ص
١٩٤٨	٥٠ر٠	صفر	صفر	صفر
١٩٤٩	٣٦ر٥	١	١	٣٦ر٥
١٩٥٠	٤٣ر٠	٢	٤	٨٦ر٠
١٩٥١	٤٤ر٥	٣	٩	١٣٣ر٥
١٩٥٢	٣٨ر٩	٤	١٦	١٥٥ر٦
١٩٥٣	٣٨ر١	٥	٢٥	١٩٠ر٥
١٩٥٤	٣٢ر٦	٦	٣٦	١٩٥ر٦
١٩٥٥	٣٨ر٧	٧	٤٩	٢٧٠ر٩
١٩٥٦	٤١ر٧	٨	٦٤	٣٣٣ر٦
١٩٥٧	٤١ر١	٩	٨١	٣٦٩ر٩
١٩٥٨	٣٨ر٨	١٠	١٠٠	٣٨٨ر٠
المجموع	٤٣٨ر٩	٥٥	٣٨٥	٢١١٠ر١

مح ص = م مح س + ن ب   ٤٣٨ر٩ = ٥٥ م + ١١ ب   (١)

محس ص = م مح س^٢ + ب مح س   ٢١١٠ر١ = ٣٨٥ م + ٥٥ ب   (٢)

بضرب المعادلة (١) في ٥   ٢١٩٤ر٥ = ٢٧٥ م + ٥٥ ب

٢١١٠ر١ = ٣٨٥ م + ٥٥ ب

بالطرح   ٨٤ر٤ = ١١٠ م

∴ م = ٨٤ر٤ / ١١٠٠ = − ٠ر٧٦٧

أي ان الاتجاه العام لهـذه السلسلة هبوطي ويساوي ٧٦٧ر٠ في العام الواحد .

بالتعويض في المعادلة (١) ٩ر٤٣٨ = – ١٨٥ر٤٢ + ١١ ب

∴ ٠٨٥ر٤٨١ = ١١ ب

∴ ب = ٧٣٥ر٤٣

∴ معادلة الاتجاه العام من نقطة الاصل عام ١٩٤٨

ص = – ٧٦٧ر٠ س + ٧٣٥ر٤٣

مثال ٧١ :

ويمكن أخذ عام ١٩٥٣ كنقطة أصل وبذلك يكون لدينا كالآتي :

السنوات	القيم (ص)	س	س ص	$س^2$
١٩٤٨	٠ر٥٠	– ٥	– ٢٥٠	٢٥
١٩٤٩	٥ر٣٦	– ٤	– ١٤٦	١٦
١٩٥٠	٠ر٤٣	– ٣	– ١٢٩	٩
١٩٥١	٥ر٤٤	– ٢	– ٨٩	٤
١٩٥٢	٩ر٣٨	– ١	– ٩ر٣٨	١
١٩٥٣	١ر٣٨	صفر	صفر	صفر
١٩٥٤	٦ر٢٢	+ ١	+ ٦ر٣٢	١
١٩٥٥	٧ر٣٨	+ ٢	+ ٤ر٧٧	٤
١٩٥٦	٧ر٤١	+ ٣	+ ١ر١٢٥	٩
١٩٥٧	١ر٤١	+ ٤	+ ٤ر١٦٤	١٦
١٩٥٨	٨ر٣٣	+ ٥	+ ٠ر١٦٩	٢٥
المجموع	٩ر٤٣٨	صفر	– ٤ر٧٤	١١٠

٤٣٨ر٩ = ١١ ب ∴ ب = ٣٩ر٩

– ٨٤ر٤ = ١١٠ م ∴ م = – ٠ر٧٦٧

∴ معادلة الاتجاه العام من نقطة الاصل عام ١٩٥٣

ص = – ٠ر٧٦٨ س + ٣٩ر٩

(نلاحظ أن قيمة م لم تتغير حيث ان ميل خط الاتجاه العام لا يمكن أن يتغير ، أما قيمة ب فتغيرت حيث انها تمثل الجزء المقطوع من المحور الرأسي بواسطة خط الاتجاه العام وهذا الجزء لا بد ان يتغير إذا غيرنا نقطة الأصل الخاصة بالسلسلة) .

وعلى أساس معادلة الاتجاه العام يمكن حساب القيم الاتجاهية أي قيم الظاهرة وهي متأثرة فقط بالاتجاه العام وذلك بالتعويض في معادلة الاتجاه عن س حسب ترتيب السنوات .

فإذا أردنا حساب القيمة الاتجاهية لعام ١٩٤٨ يكون العمل كالآتي :

ص = – ٠ر٧٦٧ × – ٥ + ٣٩ر٩
= ٣ر٨٣٥ + ٣٩ر٩
= ٤٣ر٧٣٥

وسواء أخذنا هذه المعادلة أو المعادلة السابقة فاننا نحصل على نفس النتيجة للقيم الاتجاهية :

ص = – ٠ر٧٦٧ × صفر + ٤٣ر٧٣٥
= ٤٣ر٧٣٥

وتكون القيم الاتجاهية لجميع السنوات كالآتي : ٤٣ر٧ ، ٤٣ر٠ ، ٤٢ر٢ ، ٤١ر٤ ، ٤٠ر٧ ، ٣٩ر٩ ، ٣٩ر١ ، ٣٨ر٤ ، ٣٧ر٦ ، ٣٦ر٨ ، ٣٦ر١ .

وإذا كان عدد السنوات زوجياً فلا نستطيع أن نختصر العمل كثيراً بأخذ نقطة الأصل في الوسط حيث لا توجد سنة متوسطة يمكن أن تعتبر نقطة الأصل ، ولذلك تؤخذ نقطـــة الأصل في منتصف الفئة بين السنتين المتوسطتين وبذلك يكون وحدات الزمن هي نصف السنة ويتضح ذلك من المثال الآتي :

مثال ٧٢ :

السنوات	القيم (ص)	س	س^{٢}	س ص
١٩٤٨	٥٠ر٠	٩ —	٨١	٤٥٠ر٠ —
١٩٤٩	٣٦ر٥	٧ —	٤٩	٢٥٥ر٥ —
١٩٥٠	٤٣ر٠	٥ —	٢٥	٢١٥ر٠ —
١٩٥١	٤٤ر٥	٣ —	٩	١٣٣ر٥ —
١٩٥٢	٣٨ر٩	١ —	١	٣٨ر٩ —
		صفر	صفر	صفر
١٩٥٣	٣٨ر١	١	١	٣٨ر١
١٩٥٤	٣٢ر٦	٣	٩	٩٧ر٨
١٩٥٥	٣٨ر٧	٥	٢٥	١٩٣ر٥
١٩٥٦	٤١ر٧	٧	٤٩	٢٩١ر٩
١٩٥٧	٤١ر١	٩	٨١	٣٦٩ر٩
المجموع	٤٠٥ر١	صفر	٣٣٠	١٠١ر٧ —

∴ ٤٠٥ر١ = ١٠ ب ∴ ب = ٤٠ر٥١

— ١٠١ر٧ = ٣٣٠ م ∴ م = ٠ر٣٠٨

وبذلك يكون الاتجاه العام النصف سنوي هبوطي ويساوي ٣٠٨ر٠ في المتوسط فاذا أردنا حساب القيمة الاتجاهية لعام ١٩٤٨ يكون العمل كالآتي : –

ص = − ٣٠٨ × - ٩ + ٤٠ر٥١

= ٢ر٧٧٢ + ٤٠ر٥١

= ٤٣ر٢٨٢

وعلى ساس معادلة الاتجاه العام يمكن للتنبؤ بمـــا ستكون عليه قيمة الظاهرة في المستقبل على فرض ثبات العوامل المختلفة المؤثرة على هذه الظاهرة فإذا أردنا تقديركم ستكون عليه قيمة الظاهرة عام ١٩٥٨ واستخدمنا المعادلة الأخيرة التي تكون فيها س ممثلة لوحدات نصف سنوية يكون العمل كالآتي :

ص = − ٠ر٣٠٨ × ١١ ÷ ٤٠ر٥١

= ٣ر٣٨٨ + ٤٠ر٥١

= ٣٧ر١٣٢

ولا شك أن توفيق خط الاتجاه العام بطريقة المربعات الصغرى ادق بكثير من طريقة المتوسطات المتحركة . على أن دقـــة نتائجها تتوقف على تحديد نوع الخط الذي يناسب تغير الظاهرة في المـــدة موضوع البحث ، وبشكل آخر تتوقف على تحديد المعـادلة التي تصاح لتمثيل الاتجاه العام للظاهرة .

حساب التغيرات الموسمية :

حيث أن التغيرات الموسمية تحدث في أجزاء السنة ، لذلك يجب أن تكون قيم الظاهرة على أساس الأجزاء التي نريد حساب التغيرات الموسمية الخاصة بها ، فاذا كنا نريد حساب التغيرات الموسمية في أشهر السنة (مثلا)

يجب أن تكون قيم الظاهرة متوفرة على أساس شهري ، وإذ كنا نريد حساب التغيرات الموسمية الربع سنوية يجب أن تكون القيم ربع سنوية...الخ.

وهناك عدة طرق لحساب التغيرات الموسمية غير أننا سنكتفي بشرح طريقتين منها .

١ – ايجاد النسب الموسمية بطريقة المتوسطات البسيطة .

تعتمد هذه الطريقة على حساب المتوسط العام لكل فترة خلال المدة موضوع البحث ثم حساب نسبة قيمة الظاهرة في كل فترة الى هذا المتوسط العام فاذا زادت النسبة عن ١٠٠ يدل ذلك على وجود تغير موسمي صعودي في هذه الفترة وبالعكس اذا نقصت النسبة عن ١٠٠ يدل ذلك على وجود تغير موسمي هبوطي :

مثال ٧٣ :

السنوات	يناير	فبراير	مارس	ابريل	مايو	يونيو	يوليو	أغسطس	سبتمبر	اكتوبر	نوفمبر	ديسمبر
١٩٥١	٣١٨	٢٨١	٢٧٨	٢٥٠	٢٣١	٢١٦	٢٢٣	٢٤٥	٢٦٩	٣٠٢	٣٢٥	٣٤٧
١٩٥٢	٣٤٢	٣٠٩	٢٩٩	٢٦٨	٢٤٩	٢٣٦	٢٤٢	٢٦٢	٢٨٨	٣٢١	٣٤٢	٣٦٤
١٩٥٣	٣٦٧	٣٢٨	٣٢٠	٢٨٧	٢٦٩	٢٥١	٢٥٩	٢٨٤	٣٠٩	٣٤٥	٣٦٧	٣٩٤
١٩٥٤	٣٩٢	٣٤٩	٣٤٢	٣١١	٢٩٠	٢٧٣	٢٨٢	٣٠٥	٣٢٨	٣٦٤	٣٨٩	٤١٧
١٩٥٥	٤٢٠	٣٧٨	٣٧٠	٣٣٤	، ١٤	٢٩٦	٣٠٥	٣٣٠	٣٥٦	٣٩٦	٤٤٢	٤٥٢
١٩٥٦	٤٥٣	٤١٢	٣٩٨	٣٦٢	٣٤١	٣٢٢	٢٢٥	٣٥٩	٣٩٢	٤٢٧	٤٥٤	٤٨٣
١٩٥١	٣٨٧	٤٤٠	٤٢٩	٣٩٣	٣٧٠	٣٤٧	٣٥٧	٣٨٨	٤١٥	٤٥٧	٤٩١	٥١٦
١٩٥٨	٥٢٩	٤٧٧	٤٦٣	٤٢٣	٣٩٨	٣٨٠	٣٨٩	٤١٩	٤٤٨	٤٩٣	٥٢٦	٥٦٠

المجموع الشهري :

٣٣٠٨ ٢٩٧٤ ٢٨٩٩ ٢٦٢٨ ٢٤٦٢ ٢٣٢١ ٢٣٨٢ ٢٥٩٢ ٢٨٠٥ ٣١٠٥
٣٣١٦ ٣٥٣٣ .

المتوسط الشهري

٤١٣ ٣٧٢ ٣٦٢ ٣٢٨ ٣٠٧ ٢٩٠ ٢٩٩ ٣٢٤ ٣٥٠ ٣٨٨ ٤١٤ ٤٤١ .

وحيث ان المتوسط الشهري العام = ٣٥٧ .

∴ النسب الموسمية الشهرية هي :

$\frac{٤١٣}{٣٥٧}$ × ١٠٠ = ١١٦ تقريباً أي صعود موسمي بنسبة ١٦٪

$\frac{٣٧٢}{٣٥٧}$ × ١٠٠ = ١٠٤ تقريباً أي صعود موسمي بنسبة ٤٪

$\frac{٣٦٢}{٣٥٧}$ × ١٠٠ = ١٠١ تقريباً أي صعود موسمي بنسبة ١٪

$\frac{٣٢٨}{٣٥٧}$ × ١٠٠ = ٩١ تقريباً أي هبوط موسمي بنسبة ٩٪

وهكذا بالنسبة للأشهر الباقية .

وإذا كانت قيم الظاهرة متأثرة باتجاه عام معين فلا بد من استبعاد هذ الاتجاه قبل حساب النسب الموسمية حيث اننا إذا لم نستبعده فان التغيرات الموسمية لا تكون دقيقة ، فاذا كان الاتجاه العام صعودياً ، وكان التغير الموسمي صعودياً كذلك فان النسب الموسمية تكون أكثر مما يجب أن تكون والعكس إذا كان الاتجاه العام صعودياً وكان التغير الموسمي هبوطياً فان

النسب الموسمية تكون أقل مما يجب أن تكون ، ولاستبعاد الاتجاه العام نوجد معادلة الاتجاه العام للقيم الموجودة لدينا ومنهـا نحسب القيم الاتجاهية للظاهرة ثم نقسم كل قيمة واقعية على القيمة الاتجاهيـــة المقابلة لها (نسبة مئوية) نحصل على نسب تدل على الظاهرة مستبعداً منها الاتجاه العام .

مثال ٧٤ :

(من التمرين السابق)

السنوات	المتوسط الشهري	س	س ص	س²
١٩٥١	٢٧٣ر٧	٧ -	١٩١٥ر٩ -	٤٩
١٩٥٢	٢٩٣ر٥	٥ -	١٤٦٧ر٥ -	٢٥
١٩٥٣	٣١٥ر٠	٣ -	٩٤٥ر٠ -	٩
١٩٥٤	٣٣٦ر٨	١ -	٣٣٦ر٨ -	١
١٩٥٥	٢٦٤ر٤	١	١٦٤ر٤	١
١٩٥٦	٣٩٤ر٨	٣	١١٨٤ر٤	٩
١٩٥٧	٤٢٤ر٢	٥	٢١٢١ر٠	٢٥
١٩٥٨	٤٥٨ر٧	٧	٣٢١٠ر٩	٤٩
المجموع	٢٨٦١ر١	صفر	٢٢١٥ر٥	١٦٨

المتوسط الشهري = $\frac{\text{مجموع القيم في العام}}{١٢}$

مجـ ص = م مجـ س + ن ب   ٢٨٦١ر١ = ٨ ب

مجـ س ص = م مجـ س + ب مجـ س   ٢٢١٥ر٥ = ١٦٨ م

$$\therefore \text{ب} = \frac{١ر٢٨٦١}{٨} = ٦ر٣٥٧$$

$$\text{م} = \frac{٥ر٢٢١٥}{١٦٨} = ١٨٨ر١٣$$

∴ معادلة الاتجاه العام = ١٨٨ر١٣ س + ٦ر٣٥٧

حيث س تمثل وحدات نصف سنوية من نقطة الأصل ١ يناير ١٩٥٥ ،

$$\text{وبذلك يكون الاتجاه العام الشهري} = \frac{١٨٨ر١٣}{٦} = ٢ر٢ \text{ تقريباً .}$$

وعلى أساس أن القيم الشهرية في التعرين السابق هي على أساس منتصف الشهر تكون القيمة الاتجاهية لمنتصف شهر يناير ١٩٥١ =

= ٢ر٢ × − ٥ر٤٧ + ٦ر٣٥٧

= − ٥ر١٠٤ + ٦ر٣٥٧

= ١ر٢٥٣

والقيمة الاتجاهيه لمنتصف شهرفبراير ١٩٥١ = ٢ر٢ ×− ٥ر٤٦ + ٦ر٣٥٧

= ٣ر٢٥٥

والقيمةالاتجاهية لمنتصف شهر مارس ١٩٥١ =٢ر٢ × −٥ر٤٥ + ٦ر٣٥٧

= ٥ر٢٥٧

والقيمة الاتجاهية لمنتصف شهر يناير ١٩٥٥ = ٢ر٢ × ٥ر٠ + ٦ر٣٥٧

= ٧ر٣٥٨

والقيمة الاتجاهية لمنتصف شهر فبراير ١٩٥٥ = ٢ر٢ × ٥ر١ + ٦ر٣٥٧

٩ر٣٦٠

وهكذا بالنسبة للاشهر الباقية(نلاحظ ان نقطة الأصل هو ١ يناير١٩٥٥

وبذلك يكون ترتيب منتصف شهر يناير ١٩٥١ = ٥ر٤٧، وترتيب منتصف شهر فبراير ١٩٥١ = ٥ر٤٦ وهكذا ...) . وتكون بذلك القيم الاتجاهية لجميع الأشهر موضوع البحث كالآتي :

السنوات	يناير	فبراير	مارس	ابريل	مايو	يونيو	يوليو	اغسطس	سبتمبر	اكتوبر	نوفبر	ديسمبر
١٩٥١	٢٥٣	٢٥٥	٢٥٧	٢٦٠	٢٦٢	٢٦٤	٢٦٦	٢٦٨	٢٧٠	٢٧٣	٢٧٥	٢٧٧
١٩٥٢	٢٧٩	٢٨٢	٢٨٤	٢٨٦	٢٨٨	٢٩٠	٢٩٣	٢٩٥	٢٩٧	٢٩٩	٣٠١	٣٠٣
١٩٥٣	٣٠٦	٣٠٨	٣١٠	٣١٢	٣١٥	٣١٧	٣١٩	٣٢١	٣٢٣	٣٢٦	٣٢٨	٣٣٠
١٩٥٤	٢٣٢	٤٣٤	٣٣٧	٣٣٩	٣٤١	٣٤٣	٣٤٥	٣٤٧	٣٥٠	٣٥٢	٣٥٤	٣٥٦
١٩٥٥	٣٥٩	٣٦١	٣٦٣	٣٦٥	٢٦٧	٣٧٠	٣٧٢	٢٧٤	٣٧٦	٣٧٨	٣٨١	٣٨٣
١٩٥٦	٣٨٥	٣٨٧	٣٨٩	٣٩٢	٣٩٤	٣٩٦	٣٩٨	٤٠٠	٤٠٣	٤٠٥	٤٠٧	٤٠٩
١٩٥٧	٤١١	٤١٤	٤١٦	٤١٨	٤٢٠	٤٢٢	٤٢٥	٤٢٧	٤٢٩	٤٣١	٤٣٣	٤٣٦
١٩٥٨	٤٣٨	٤٤٠	٤٤٢	٤٤٤	٤٤٧	٤٤٩	٤٥١	٤٥٣	٤٥٥	٤٥٨	٤٦٠	٤٦٢

وبقسمة القيم الشهرية الأصلية على القيم الاتجاهية نحصل على النسب الآتية :

السنوات	ينايو	فبراير	مارس	ابريل	مايو	يونيو	يوليو	اغسطس	سبتمبر	!اكتوبر	نوفبر	ديسمبر
١٩٥١	١٢٦	١١٠	١٠٨	٩٦	٨٨	٨٢	٨٤	٩١	٩٩	١١١	١١٨	١٢٥
١٩٥٢	١٢٢	١١٠	١٠٥	٩٤	٨٦	٨١	٨٣	٨٩	٩٧	١٠٧	١١٣	١٢٠
١٩٥٣	١٢٠	١٠٦	١٠٣	٩٢	٨٥	٧٩	٨١	٨٨	٩٥	١٠٦	١١٢	١١٩
١٩٥٤	١١٨	١٠٤	١٠٢	٩٢	٨٥	٧٩	٨٢	٨٨	٩٤	١٠٣	١١٠	١١٧
١٩٥٥	١١٧	١٠٥	١٠٢	٩١	٨٥	٨٠	٨٢	٨٨	٩٥	١٠٥	١١١	١١٨
١٩٥٦	١١٨	١٠٦	١٠٢	٩٢	٨٧	٨١	٨٤	٩٠	٩٧	١٠٥	١١٢	١١٨
١٩٥٧	١١٨	١٠٦	١٠٣	٩٤	٨٨	٨٢	٨٤	٩١	٩٧	١٠٦	١١٣	١١٨
١٩٥٨	١٢١	١٠٨	١٠٥	٩٥	٨٩	٨٥	٨٦	٩٢	٩٨	١٠٨	١١٤	١٢١

وبذلك تكون المتوسطات الشهرية هي كالآتي :

١٢٠ ، ١٠٧ ، ١٠٤ ، ٩٣ ، ٨٧ ، ٨١ ، ٨٤ ، ٩٠ ، ٩٦ ، ١٠٦ ، ١١٣ ، ١١٩ ، وحيث ان مجموع هذه النسب = ١٢٠٠ فتكون النسب السابقة هي نفسها تعبر عن النسب الموسمية ، أما إذا كان مجموعها ليس ١٢٠٠ فيجب تصحيحها حتى نحصل على النسب الموسمية ، فاذا فرضنا أن مجموعها = ١١٠٠ ، فان النسبة الاولى تصبح بعد التصحيح ١٢٠ × $\frac{١٢٠٠}{١١٠٠}$ = ١٣١ تقريباً ، وهكذا بالنسبة لباقي المتوسطات الشهرية .

٢ – ايجاد التقلبات الموسمية بوحدات مطلقة .

قد نحتاج أحياناً إلى حساب القيم الفعلية للتقلبات الموسميه . ولحساب هذه القيم نتبع الخطوات الآتيه :

١ – نحسب القيم الاتجاهية سواء على اساس المتوسطات المتحركة أو على أساس معادلة الاتجاه العام .

٢ – نحسب انحرافات القيم الفعلية عن القيم الاتجاهية فيكون الفرق هو عبارة عن التغيرات الموسمية والعرضية والدورية ، فاذا كنا قد أخذنا فترة ليس فيها تغيرات عرضية أو دورية يكون الفرق هو فقط التغيرات الموسمية ، أما إذا كان هناك تغيرات عرضية نحسب متوسط الانحرافات ، وبذلك نقضي على هذه التغيرات حيث أن التقلبات العرضية لا تتبع قاعدة معينة فأحياناً تكون صعودية وأحياناً تكون هبوطية ، وبذلك يتلاشى أثرها عند حساب المتوسط (هذا على أساس أن نكون قد أخذنا مدة طويلة بحيث يمكن أن تكون هناك تغيرات عرضية صعودية واخرى هبوطية) .

مثـــال ٧٥ :

فترات الربع سنوية	القيم	متوسط متحرك على اساس اربع سنوات	متوسط متحرك على اساس فترتين	انحراف القيمة الواقعية عن القيمة الاتجاهية
١٩٥١ الربع الاول	٦٥			
« الثاني	٦٢			
		٦٣	٦٢ر٥	− ١ر٥
« الثالث	٦١			
		٦٢	٦١ر٥	+ ١ر٥
« ~~الرابع~~	~~٦٣~~			
		٦١	٦٠ر٥	+ ٣ر٥
١٩٥٢ « الاول	٦٤			
		٦٠	٥٩ر٥	− ٢ر٥
« الثاني	٥٧			
		٥٩	٥٩ر٥	− ٣ر٥
« الثالث	٥٦			
		٦٠	٦٠ر٥	+ ٠ر٥
« الرابع	٦١			
		٦١	٦٢	+ ٦ر٠
١٩٥٣ « الاول	٦٨			
		٦٣	٦٣ر٥	− ١ر٥
« الثاني	٦٢			
		٦٤	٦٤ر٥	− ٣ر٥
« الثالث	٦١			
		٦٥	٦٤ر٥	+ ٢ر٥
« الرابع	٦٧			
		٦٤	٦٣ر٥	+ ٦ر٥
١٩٥٤ « الاول	٧٠			
		٦٣	٦٢	− ٣ر٠
« الثاني	٥٩			
		٦١	٦٠	− ٤ر٠
« الثالث	٥٦			
		٥٩	٥٧	+ ٣ر٠
« الرابع	٦١			
		٥٧	٥٦ر٥	+ ٢ر٥
١٩٥٥ « الاول	٥٩			
		٥٦	٥٥ر٥	− ١ر٥
« الثاني	٥٤			
		٥٥		
« الثالث	٥١			
« الرابع	٥٨			

فاذا اخذنا الربع الاول من كل عام وحسبنا متوسط الانحرافات يكون الناتج هو التغير الموسمي ، + ٣ر٥ ، + ٦ + ٦ر٥ ، + ٢ر٥ ومتوسطها + ٤ر٦ وهو قيمة التغير الموسمي في هذا الربع – وفي الربع الثاني – ٢ر٥ ، – ١ر٥ ، – ٣ ، – ١ر٥ ومتوسطها – ٢ر١ وهو قيمة التغير الموسمي في هذا الربع . وفي الربع الثالث – ١ر٥ ، – ٣ر٥ ، – ٣ر٥ ، – ٤ ومتوسطها – ٢ر٨٧٥ وهو قيمة التغير الموسمي في هذا الربع . وفي الربع الرابع ١ر٥ ، ٠ر٥ ، ٢ر٥ ، ٣ ومتوسطها + ١ر٨٧٥ وهوقيمة التغير الموسمي في هذا الربع.

استبعاد تأثير التقلبات الموسمية :

من الواضح ان الهدف من دراسة التقلبات الموسمية هو تحديد تأثير كل موسم على قيم الظاهرة موضوع البحث حتى يمكن الاستعداد لمواجهة الزيادة الموسمية او النقص الموسمي ولهذا اهمية كبيرة بالنسبة للمنتجين او التجار الذين تتعرض منتجاتهم لتقلبات موسمية . ومن ناحية اخرى فان المنتج يهمه كذلك ان يتعرف على الحالة الحقيقية للطلب على منتجاته ، اي حالة الطلب بدون التأثيرات الموسمية ، المختلفة حتى لا تضلله الزيادة المؤقتة او النقص المؤقت في مبيعاته ، ويمكن استبعاد تأثير الموسم بقسمة القيمة الواقعية للظاهرة في الفترة موضوع البحث على النسبة الموسمية في هذة الفترة . فاذا فرضنا مثلا ان قيمة مبيعات محل تجاري في شهر معين هي ٥٠٠٠٠ ليرة وكانت النسبة الموسميه للمبيعات في هذا الشهر هي ١٢٥ ٪ تكون مبيعات المحل بدون تأثير الموسم أي بعد استبعاده = $٥٠٠٠٠ \times \frac{١٠٠}{١٢٥}$ ٤٠٠٠٠ ليرة . أما اذا كانت النسبة الموسمية تدل على هبوط مثلا ٨٠ ٪ تكون قيمة المبيعات بدون تأثير الموسم = $٥٠٠٠٠ \times \frac{١٠٠}{٨٠}$ = ٦٢٥٠٠ واذا كانت التغيرات الموسمية معروفة

في شكل قيم مطلقة يكون استبعادها بطرحها مباشرة من القيم الحقيقية للظاهرة (طرحا جبريا) .

التغيرات العرضية والدورية :

اذا توفرت لدينا القيم الواقعية والاتجاهية لظاهرة ما في فترة زمنية معينة فان الفرق بينهما يدلنا على التغيرات الموسمية والعرضية ، فاذا طرحنا قيمة التغير الموسمي يكون الباقي هو التغير العرضي في هذه الفترة .

مثـــال ٧٦ :

اذا كانت القيمة الواقعية = ٦٤ والقيمـــة الاتجاهية = ٦٥ والتغير الموسمي = − ٣

التغير العرضي = ٦٤ − (٦٥ + (− ٣))

= ٦٤ − ٦٢ = ٢

وإذا كانت القيمة الواقعيـــة = ٦٦ والقيمة الاتجاهية = ٦٢ والتغير الموسمي = + ٤ر٤

∴ التغير العرضي = ٦٦ − (٦٢ + (+ ٤ر٤))

= ٦٦ − ٦٦ر٤

= [illegible]

وإذا لم تكن الظاهرة قد أصيبت بتغيرات عرضية يكون الفرق بين القيمة الواقعية والقيمة الاتجاهيـــة هو التغيرات الموسمية والتغيرات الدورية ، فاذا طرحنا قيمة التغير الموسمي يكون الباقي هـو التغير الدوري . كذلك يمكن بقسمة القيمة الواقعية للظاهرة على قيمتها الاتجاهية أن نحصل على نسبة مئوية تدل على الظاهرة بدون الاتجاه العام ، وإذا قسمنا هذه النسبة على النسبة

الموسمية في الفترة موضوع البحث نحصل على نسبة مئوية تدل على التغير الدوري والعرضي سويـــاً ، فاذا كانت لدينا فترة طويلة يمكننا أن نستبعد التغيرات العرضية بحساب المتوسطات المتحركة للنسب التي تحصل عليها فتكون النسب بعد استبعاد التغيرات العرضية معبرة عن التغيرات الدورية . وبذلك نستطيع أن نستنتج القاعدة بأن القيمة المشاهدة (الواقعيه) = القيمة الاتجاهية × التغير الموسمي × التغير الدوري × التغير العرضي .

مثال ٧٠ :

الأشهر	القيم الواقعية	القيم الاتجاهية	القيمة الواقعية / القيمة الاتجاهية × ١٠٠	النسب الموسمية	النسب الدورية
ينـاير	١٦٩	٣٩٥	٤٣ ٪	٧٥ ٪	٥٧ ٪
فبراير	٢١٢	٣٩٦	٥٤	٩٦	٥٦
مارس	٢٥١	٣٩٧	٦٣	١٠٧	٥٩
ابريل	٢٨٦	٣٩٨	٧٢	١٢٣	٥٨
مـايو	٢٨٥	٣٩٩	٧١	١٢٢	٥٨
يونيو	٢٦٤	٤٠٠	٦٦	١١١	٥٩
يونيو	١٦٤	٤٠١	٤٩	٨٣	٤٩
اغسطس	١٩٦	٤٠٢	٤٩	٧٢	٦٨
سبتمبر	٢٠٨	٤٠٣	٥٢	٨٨	٥٩
اكتوبر	٢٦٤	٤٠٤	٦٥	١٠٨	٦٠
نوفمبر	٢٧٤	٤٠٥	٦٨	١١٠	٦٢
ديسمبر	٢٧٣	٤٠٦	٦٧	١٠٣	٦٥

نسب التغيرات الدورية = النسبة بدون الاتجاه العام ÷ النسبة الموسمية .
وتطبيقاً للقاعدة التي سبق الاشارة اليهـــا – القيمة المشاهدة = القيمة الاتجاهية × التغير الموسمي × التغير الدوري × التغير العرضي يمكن التنبؤ بما سوف تكون قيمة الظاهرة في فترة مستقبلة . ولما كان التغير العرضي لا يمكــن التنبؤ به لهذا يكون تنبؤنا قاصراً على قيمة الظاهرة وهي متأثرة فقط بالاتجاه العام والتغير الموسمي والتغير الدوري .

مثال ٧٨ :

اذا كانت القيمة الاتجاهية المتوقعة في شهر ما هي ٣٩٥ وان النسبة الموسمية لهذا الشهر هي ٧٥٪ والنسبة الدورية ٥٧٪ .

∴ قيمة الظاهرة وهي متأثرة بهذه الأنواع الثلاث من التغيرات .

$$= ٣٩٥ \times \frac{٧٥}{١٠٠} \times \frac{٥٧}{١٠٠} = ١٦٨٫٨٦$$ أي ١٦٩ تقريباً .

الارتباط بين السلاسل الزمنية :

تبين لنا مما تقدم ان السلاسل الزمنية تتأثر باربع أنواع من التغيرات – الاتجاه العام والتقلبات الموسمية والدورية والعرضية . وعندما نريد دراسة الارتباط بين سلسلتين زمنيتين يجب أن نحدد نوع العلاقة التي نريد دراستها ، فقد ترغب في بحث ما اذا كان هناك ارتباط بين الاتجاه العام لسلسلة مـــا والاتجاه العام لسلسلة اخرى . واذا كان هذا هو هدفنا نحسب القيم الاتجاهية لكل من السلسلتين ثم نرصد هذه القيم على ورق الرسم البياني وسوف يتضح من الرسم مباشرة اذا كان هناك اتفاق بين اتجاه الظاهرتين ، الأمر الذي يدل على وجود الارتباط بينهما ، وفي هذه الحالة لا يكون هناك ضرورة لحساب معامل الارتباط حيث أن مجرد الانفاق بين الاتجاه العا للظاهرتين يدل على

وجود ارتباط قوي مهما كانت درجة انحدار كل من خطي الاتجاه العام .

على أن الذي يهمنـــا عند دراسة الارتباط بين سلسلتين هو الاجابة على السؤال – هل هناك علاقة بين ذبذبة سلسلة ما حول اتجاهها العــام وذبذبة سلسلة أخرى حول اتجاهها العام أو لا ، وما درجة هذه العلاقة ؟

فاذا كان هذا هو هدفنا يتحتم علينا حسب قيم هذه الذبذبات في كل من السلسلتين ثم نقوم بحساب الارتباط بين هذه القيم باي قـــانون من قوانين الارتباط .

مثـــال ٧٩ :

السنوات	الرقم القيـــاسي للانتاج الصناعي	الرقم القيـــاسي لأسعار الأسهم الصناعية
١٩٥٤	١٠١	١٠٩
١٩٥٥	٦٧	٨٢
١٩٥٦	٨٥	١٢٠
١٩٥٧	١٠٥	١٠٣
١٩٥٨	٩٨	٩٠
١٩٥٩	١١٠	٩٦
١٩٦٠	١٠٨	٩٩

١ – نحسب اولا القيم الاتجاهية لكل من السلسلتين بطريقة المربعات الصغرى .

معادلة الاتجاه العام للرقم القياسي للانتاج الصناعي :

س	ص	س ص	س^2
– ٣	١٠١	– ٣٠٣	٩
– ٢	٦٧	– ١٣٤	٤
– ١	٨٥	– ٨٥	١
صفر	١٠٥	صفر	صفر
١	٩٨	٩٨	١
٢	١١٠	٢٢٠	٤
٣	١٠٨	٣٢٤	٩
المجمـــوع	٦٧٤	+ ١٢٠	٢٨

مج ص = م مج س + ن   ٦٧٤ = ٧ ب

مج س ص = م مج س^2 + ب مج س   ١٢٠ = ٢٨ م

$$\therefore \text{ب} = \frac{٦٧٤}{٧} = ٩٦٫٢٨٥٧$$

$$\text{و م} = \frac{١٢٠}{٢٨} = ٤٫٢٨٥٧$$

∴ ص = ٤٫٢٨٥٧ س + ٩٦٫٢٨٥٧

معادلة الاتجاه العام للرقم القياسي لأسعار الاسهم الصناعية

س	ص	ص س	$س^2$
٣ –	١٠٩	٣٢٧ –	٩
٢ –	٨٢	١٦٤ –	٤
١ –	١٢٠	١٢٠ –	١
صفر	١٠٣	صفر	صفر
١	٩٠	٩٠	١
٢	٩٦	١٩٢	٤
٣	٩٩	٢٩٧	٩
المجموع	٦٩٩	٣٢ –	٢٨

مجـ ص = مجـ م س + ن ب   ٦٩٩ = ٧ ب

مجـ س ص = م مجـ $س^2$ + ب مجـ س   −٣٢ = ٢٨ م

∴ ب = $\frac{٦٩٩}{٧}$ = ٩٩٫٨٥٧

و م = $\frac{٣٢}{٢٨}$ = − ١٫١٤٣

∴ ص = − ١٫١٤٣ س + ٩٩٫٨٥٧

٢ – نحسب القيم الاتجاهية بالتعويض في معادلات الاتجاه العام عن ص ثم نحسب انحرافات القيم الواقعية عن القيم الاتجاهية :

انحرافــات قيم الانتاج الصناعي	انحرافات قيم الاسهم	مربع انحرافات قيم الانتاج الصناعي	مربع انحرافات قيم الأسهم	حاصل ضرب الانحرافات
+ ١٧ر٦	+ ٥ر٧	٣٠٩ر٧٦	٣٢ر٤٩	+ ١٠٠ر٣٢
– ١٠ر٧	– ٢٠ر١	٤٢٨ر٤٩	٤٠٤ر٠١	+ ٤١٦ر٠٧
– ٧ر٠	+ ١٩ر٠	٤٩ر٠٠	٣٦١ر٠٠	– ١٣٣ر٠٠
+ ٨ر٧	+ ٣ر٤	٧٥ر٦٩	٩ر٦١	+ ٢٦ر٩٧
– ٢ر٦	– ٨ر٧	٦ر٧٦	٧٥ر٦٩	+ ٢٢ر٦٢
+ ٥ر١	– ١ر٦	٢٦ر٠١	٢ر٥٦	– ٨ر١٦
– ١ر١	+ ٢ر٦	١ر٢١	٦ر٧٦	– ٢ر٨٦
المجمــــوع	–	٨٩٦ر٩٢	٨٩٢ر١٢	+ ٩٢١ر٩٦

(نلاحظ ان مجموع الانحرافات عن القيم الاتجـــــاه = صفر لكل من الظاهرتين)

∴ الانحراف المعياري لانحرافات الرقم القياسي للانتاج الصناعي =

$$\sqrt{\frac{٨٩٦ر٩٢}{٧}} = ١١ر٣$$

والانحراف المعياري لانحرافات الرقم القياسي لأسعار الأسهم =

$$\sqrt{\frac{١٢ر٨٩٢}{٧}} = ٢ر١١$$

$$\therefore\ ر = \frac{٩٦ر٤٢١}{٧ \times ٣ر١١ \times ٢ر١١} = \frac{٩٦ر٤٢١}{٩٢ر٨٨٥} = ٤٧ر٠$$

تمـــارين

١ – الجدول الآتي يبين انتاج الصلب في احدى الدول :

السنة	الانتـــاج	السنة	الانتـــاج
١٩٥٠	٣٢١٠ طن	١٩٥٥	٥٨٣٠ طن
١٩٥١	٣٤٢٦ طن	١٩٥٦	٥٩٧٠ طن
١٩٥٢	٣٦٦٩ طن	١٩٥٧	٦٢٠٠ طن
١٩٥٣	٤٩٠٠ طن	١٩٥٨	٦٦٥٠ طن
١٩٥٤	٥٣٠٠ طن	١٩٥٩	٧٢٥٠ طن

اوجد معادلة الاتجاه العام بطريقة المربعات الصغرى ، ومنهـــا استنتج القيم الاتجاهية في الأعوام ١٩٤٥ ، ١٩٥٥ ، ١٩٦٥ .

٢ – الأرقام الآتية تبين النسب النموذجية للنقد المتداول :

يناير	٩٥	يوليو	٩٣
فبراير	٩٩	اعسطس	٩٨
مارس	١٠٠	سبتمبر	١٢٠

ابريل	٩٩	اكتوبر	١٠٩
مايو	٩٤	نوفمبر	٩٧
يونيو	٩٢	ديسمبر	٩٦

احسب النسب الموسمية .

٣ – احسب الارتباط بين الاتجاه العام للسلسلتين الآتيتين ، ثم بين تقلباتهما حول الاتجاه العام .

السنوات	الرقم القياسي للانتاج الزراعي	الرقم القياسي لاسعار الحاصلات الزراعية
١٩٥٠	١٣٦	٨٩
١٩٥١	٧٦	٦٤
١٩٥٢	١١٤	٦٥
١٩٥٣	١٣٤	٧٦
١٩٥٤	٨٣	٧٥
١٩٥٥	١٠٢	٩٢
١٩٥٦	٨٨	١١٨
١٩٥٧	٩٦	١١٥
١٩٥٨	١٢٠	١١٨

أعداد عشوائية

٣٢	٥٠	٥٣	١٦	٩٥	٩٨	٣٢	٤٨	٦٣	٩٦	٧٧	٠٢	٢٢	٤٩	١٥
٩٢	٥٢	٤٠	٥١	٨٢	٩٦	٤٠	٣٦	٤٠	٣٢	٦٤	٦٧	٣٦	١٢	٢٨
٢٧	٧٠	٢١	٣٤	٥٩	٠٣	٣٩	٤٣	٢٧	٥٠	١٢	٥٥	١١	٢٥	٣٤
٢٥	١٠	١٦	٧٠	٩١	٤٣	٥٣	٧١	٠٠	٥٤	٢٦	٣٧	٥٢	٨٢	٢٣
٣١	١١	٣٣	٤٩	٨٣	٠١	٠٦	٤٣	٤٩	١٤	٣٧	٤٢	٨١	٨٣	٦٧

٤٠	٥٤	٧٧	٧٨	١٩	٧٦	٢٥	٧٤	١٢	٢٥	٤٤	١٧	٢٦	٤١	٧٦
٣٤	١١	٦٤	٤٧	١٧	٣٤	١٥	٧٨	٢٠	٧٥	٦٤	٧٦	٤٥	٥٩	٨٥
١٨	٨٧	٣٧	٦١	٣٢	٣٧	٧٧	٨٠	٢٩	٨٠	٢٨	٤٧	٥٣	٢٠	٩٢
٩٤	٦٧	٩٢	٤٤	٧٩	٨٢	٧١	٦٥	٥١	٥١	٥٣	٢١	٠٩	٦٦	٥٨
٤٠	٧٣	٤٤	٢١	١٤	٢٠	٦٥	٩٨	١٤	٣٥	٥١	٩٦	٩١	٣٢	٢٧

٦٣	٧١	١٤	٤٨	٩١	٥٠	٢٧	٧٩	٥٢	٥٣	٣١	٣٠	٠٦	٧٠	٣١
٣٩	٧٠	١٥	١٨	٩٦	١٤	٨١	٨٤	٥٠	١١	٩٥	١٥	٠٠	٨٢	٤٠
٥٩	٩١	٢٥	١٠	٥٤	٧٧	٥٤	٢٣	٨٧	٣٢	٤٦	٠٦	٥٨	٦٥	٥٩
٠٥	٢٨	٥٣	٦٨	٥٢	٤٩	٣٧	٩٦	٩٧	١٩	٨٤	٢٤	٣٠	٦١	٦٤
٦٣	١٢	٨٢	٧٠	٣٥	٠٢	٥٤	١٠	٢٨	٧٤	٢١	٥٥	٤٢	٥٦	٤٣

٣٣	٤٦	٨٢	١٣	٠٥	١٦	١٣	٨٨	٧٠	٧٠	٧٠	٤٠	٢٦	٨٢	٠٩
٨٥	٦٠	٦٧	٥٥	١٤	٤٠	٨٩	٦٨	٦٥	٤٣	٨٣	٠٠	٤٠	٧٩	٦٨
٢٠	٥٥	٩٨	٧٦	٩٨	٢٥	١٩	٤٠	٤٧	٢١	٧٦	١٩	٩٢	٨٠	٣٦
٦١	٣٢	٥١	٦٦	٩٠	١٢	٥٣	٤٧	٢٥	٦٢	٠٣	٤٣	٢٤	٣٦	٠٧
٦٨	٧٧	٥٤	٥٣	٩٠	٥٨	١٤	٣٢	٣٩	١٧	٢٣	٩٥	١٣	٢٧	٧٢

٨١	٣٥	٠٦	٥٤	٢٠	٣٨	١١	٤٤	٣١	٢٩	٠٠	٤٤	٦٣	١١	٨٠
٧٣	٦٦	٣٥	٨٣	٠٨	٤٩	٤٠	٤١	١٥	٦٤	٠٤	٦٤	٥٥	٦١	٢٢
٧٣	٤٤	٤٥	٩٥	٣٨	٢٦	٠٦	٦١	١١	١٢	٥٤	٩٠	٣٧	٣٩	٨٥
٤٢	٧٨	٧٢	٥٤	٩٨	٥٧	٢٤	٣٠	٤٤	٩٢	٦٩	٠٥	٢٨	٤٣	٢١
٦٨	٧٢	٩٥	٨٣	٦٣	٧٤	١٥	٤٧	٣٨	٨٣	٩٢	٢٩	٧٥	٨٥	١١

٩٣	١٩	٩٢	٦٢	٨٣	٨٧	٣٣	٩٠	٢٧	٢٢	٥٥	٢٤	٥٤	٩٣	٢٨
٠٢	٥٤	٧٠	٧٧	٢٢	٩٢	٩٦	٧٩	٦٨	٨٥	٢٣	٢٢	٥٧	٥٧	٨٣
٦٢	٥١	٧٣	٦٠	٦٤	١٥	٨٠	٥٠	٤٣	٦٧	٤٨	٢٥	١٧	١٤	٣٠
٣٨	٠٦	١٨	٤٢	٥٧	٦٦	١٣	٧٨	٦٢	٧١	٥٢	٥٩	٧٦	٢٦	٤٨
٦١	٤٩	١٤	٧٧	٤٤	٩٩	٧٢	١٣	٣٥	٨٠	٧٢	٣٠	٩٥	٢٨	٢٢

جدول مساحات المنحنى المعتدل

س/ع	٠,٠٠	٠,٠١	٠,٠٢	٠,٠٣	٠,٠٤	٠,٠٥	٠,٠٦	٠,٠٧	٠,٠٨	٠,٠٩
٠,٠	٠,٠٠٠٠	٠,٠٠٤٠	٠,٠٠٨٠	٠,٠١٢٠	٠,٠١٥٩	٠,٠١٩٩	٠,٠٢٣٩	٠,٠٢٧٩	٠,٠٣١٩	٠,٠٣٥٩
٠,١	٠,٠٣٩٨	٠,٠٤٣٨	٠,٠٤٧٨	٠,٠٥١٧	٠,٠٥٥٧	٠,٠٥٩٦	٠,٠٦٣٦	٠,٠٦٧٥	٠,٠٧١٤	٠,٠٧٥٣
٠,٢	٠,٠٧٩٣	٠,٠٨٣٢	٠,٠٨٧١	٠,٠٩١٠	٠,٠٩٤٨	٠,٠٩٨٧	٠,١٠٢٦	٠,١٠٦٤	٠,١١٠٣	٠,١١٤١
٠,٣	٠,١١٧٩	٠,١٢١٧	٠,١٢٥٥	٠,١٢٩٣	٠,١٣٣١	٠,١٣٦٨	٠,١٤٠٦	٠,١٤٤٣	٠,١٤٨٠	٠,١٥١٧
٠,٤	٠,١٥٥٤	٠,١٥٩١	٠,١٦٢٨	٠,١٦٦٤	٠,١٧٠٠	٠,١٧٣٦	٠,١٧٧٢	٠,١٨٠٨	٠,١٨٤٤	٠,١٨٧٩
٠,٥	٠,١٩١٥	٠,١٩٥٠	٠,١٩٨٥	٠,٢٠١٩	٠,٢٠٥٤	٠,٢٠٨٨	٠,٢١٢٣	٠,٢١٥٧	٠,٢١٩٠	٠,٢٢٢٤
٠,٦	٠,٢٢٥٧	٠,٢٢٩١	٠,٢٣٢٤	٠,٢٣٥٧	٠,٢٣٨٩	٠,٢٤٢٢	٠,٢٤٥٤	٠,٢٤٨٦	٠,٢٥١٨	٠,٢٥٤٩
٠,٧	٠,٢٥٨٠	٠,٢٦١٢	٠,٢٦٤٢	٠,٢٦٧٣	٠,٢٧٠٤	٠,٢٧٣٤	٠,٢٧٦٤	٠,٢٧٩٤	٠,٢٨٢٣	٠,٢٨٥٢
٠,٨	٠,٢٨٨١	٠,٢٩١٠	٠,٢٩٣٩	٠,٢٩٦٧	٠,٢٩٩٥	٠,٣٠٢٣	٠,٣٠٥١	٠,٣٠٧٨	٠,٣١٠٦	٠,٣١٣٣
٠,٩	٠,٣١٥٩	٠,٣١٨٦	٠,٣٢١٢	٠,٣٢٣٨	٠,٣٢٦٤	٠,٣٢٨٩	٠,٣٣١٥	٠,٣٣٤٠	٠,٣٣٦٥	٠,٣٣٨٩
١,٠	٠,٣٤١٣	٠,٣٤٣٨	٠,٣٤٦١	٠,٣٤٨٥	٠,٣٥٠٨	٠,٣٥٣١	٠,٣٥٥٤	٠,٣٥٧٧	٠,٣٥٩٩	٠,٣٦٢١
١,١	٠,٣٦٤٣	٠,٣٦٦٥	٠,٣٦٨٦	٠,٣٧١٨	٠,٣٧٢٩	٠,٣٧٤٩	٠,٣٧٧٠	٠,٣٧٩٠	٠,٣٨١٠	٠,٣٨٣٠
١,٢	٠,٣٨٤٩	٠,٣٨٦٩	٠,٣٨٨٨	٠,٣٩٠٧	٠,٣٩٢٥	٠,٣٩٤٤	٠,٣٩٦٢	٠,٣٩٨٠	٠,٣٩٩٧	٠,٤٠١٥
١,٣	٠,٤٠٣٢	٠,٤٠٤٩	٠,٤٠٦٦	٠,٤٠٨٣	٠,٤٠٩٩	٠,٤١١٥	٠,٤١٣١	٠,٤١٤٧	٠,٤١٦٢	٠,٤١٧٧
١,٤	٠,٤١٩٢	٠,٤٢٠٧	٠,٤٢٢٢	٠,٤٢٣٦	٠,٤٢٥١	٠,٤٢٦٥	٠,٤٢٧٩	٠,٤٢٩٢	٠,٤٣٠٦	٠,٤٣١٩
١,٥	٠,٤٣٣٢	٠,٤٣٤٥	٠,٤٣٥٧	٠,٤٣٧٠	٠,٤٣٨٢	٠,٤٣٩٤	٠,٤٤٠٦	٠,٤٤١٨	٠,٤٤٣٠	٠,٤٤٤١
١,٦	٠,٤٤٥٢	٠,٤٤٦٣	٠,٤٤٧٤	٠,٤٤٨٥	٠,٤٤٩٥	٠,٤٥٠٥	٠,٤٥١٥	٠,٤٥٢٥	٠,٤٥٣٥	٠,٤٥٤٥
١,٧	٠,٤٥٥٤	٠,٤٥٦٤	٠,٤٥٧٣	٠,٤٥٨٢	٠,٤٥٩١	٠,٤٥٩٩	٠,٤٦٠٨	٠,٤٦١٦	٠,٤٦٢٥	٠,٤٦٣٣
١,٨	٠,٣٦٤١	٠,٤٦٤٩	٠,٤٦٥٦	٠,٤٦٦٤	٠,٤٦٧١	٠,٤٦٧٨	٠,٤٦٨٦	٠,٤٦٩٣	٠,٤٦٩٩	٠,٤٧٠٦
١,٩	٠,٤٧١٣	٠,٤٧١٩	٠,٤٧٢٦	٠,٤٧٣٢	٠,٤٧٣٨	٠,٤٧٤٤	٠,٤٧٥٠	٠,٤٧٥٨	٠,٤٧٦٢	٠,٤٧٦٧
٢,٠	٠,٤٧٧٣	٠,٤٧٧٨	٠,٤٧٨٣	٠,٤٧٨٨	٠,٤٧٩٣	٠,٤٧٩٨	٠,٤٨٠٣	٠,٤٨٠٨	٠,٤٨١٢	٠,٤٨١٧
٢,١	٠,٤٨٢١	٠,٤٨٢٦	٠,٤٨٣٠	٠,٤٨٣٤	٠,٤٨٣٨	٠,٤٨٤٢	٠,٤٨٤٦	٠,٤٨٥٠	٠,٤٨٥٤	٠,٤٨٥٧
٢,٢	٠,٤٨٦١	٠,٤٨٦٥	٠,٤٨٦٨	٠,٤٨٧١	٠,٤٨٧٥	٠,٤٨٧٨	٠,٤٨٨١	٠,٤٨٨٤	٠,٤٨٨٧	٠,٤٨٩٠
٢,٣	٠,٤٨٩٣	٠,٤٨٩٦	٠,٤٨٩٨	٠,٤٩٠١	٠,٤٩٠٤	٠,٤٩٠٦	٠,٤٩٠٩	٠,٤٩١١	٠,٤٩١٣	٠,٤٩١٦
٢,٤	٠,٤٩١٨	٠,٤٩٢٠	٠,٤٩٢٢	٠,٤٩٢٥	٠,٤٩٢٧	٠,٤٩٢٩	٠,٤٩٣١	٠,٤٩٣٢	٠,٤٩٣٤	٠,٤٩٣٦
٢,٥	٠,٤٩٣٨	٠,٤٩٤٠	٠,٤٩٤١	٠,٤٩٤٣	٠,٤٩٤٥	٠,٤٩٤٦	٠,٤٩٤٨	٠,٤٩٤٩	٠,٤٩٥١	٠,٤٩٥٢
٢,٦	٠,٤٩٥٣	٠,٤٩٥٥	٠,٤٩٥٦	٠,٤٩٥٧	٠,٤٩٥٩	٠,٤٩٦٠	٠,٤٩٦١	٠,٤٩٦٢	٠,٤٩٦٣	٠,٤٩٦٤
٢,٧	٠,٤٩٦٥	٠,٤٩٦٦	٠,٤٩٦٧	٠,٤٩٦٨	٠,٤٩٦٩	٠,٤٩٧٠	٠,٤٩٧١	٠,٤٩٧٢	٠,٤٩٧٣	٠,٤٩٧٤
٢,٨	٠,٤٩٧٤	٠,٤٩٧٥	٠,٤٩٧٦	٠,٤٩٧٧	٠,٤٩٧٧	٠,٤٩٧٨	٠,٤٩٧٩	٠,٤٩٨٠	٠,٤٩٨٠	٠,٤٩٨١
٢,٩	٠,٤٩٨١	٠,٤٩٨٢	٠,٤٩٨٣	٠,٤٩٨٤	٠,٤٩٨٤	٠,٤٩٨٤	٠,٤٩٨٥	٠,٤٩٨٥	٠,٤٩٨٦	٠,٤٩٨٦
٣,٠	٠,٤٩٨٦٥	٠,٤٩٨٧	٠,٤٩٨٧	٠,٤٩٨٨	٠,٤٩٨٨	٠,٤٩٨٨	٠,٤٩٨٩	٠,٤٩٨٩	٠,٤٩٨٩	٠,٤٩٩٠

جدول x^2

درجات الحرية	احتمال الحصول على قيمة x^2 المبينة بالجدول بطريق المصادفة						
	٠,٠٠١	٠,٠١	٠,٠٥	٠,١٠	٠,٥٠	٠,٩٠	٠,٩٩
١	١٠,٨٢٧	٦,٦٣٥	٣,٨٤١	٢,٧٠٦	٠,٤٥٥	٠,٠١٥٨	٠,٠٠٠١٥٧
٢	١٣,٨١٥	٩,٢١٠	٥,٩٩١	٤,٦٠٥	١,٣٨٦	٠,٢١١	٠,٠٢٠١
٣	١٦,٢٦٨	١١,٣٤١	٧,٨١٥	٦,٢٥١	٢,٣٦٦	٠,٥٨٤	٠,١١٥
٤	١٨,٤٦٥	١٣,٢٧٧	٩,٤٨٨	٧,٧٧٩	٣,٣٥٧	١,٠٦٤	٠,٢٩٧
٥	٢٠,٥١٧	١٥,٠٨٦	١١,٠٧٠	٩,٢٣٦	٤,٣٥١	١,٦١٠	٠,٥٥٤
٦	٢٢,٤٥٧	١٦,٨١٢	١٢,٥٩٢	١٠,٦٤٥	٥,٣٤٨	٢,٢٠٤	٠,٨٧٢
٧	٢٤,٣٢	١٨,٤٧٥	١٤,٠٦٧	١٢,٠١٧	٦,٣٤٦	٢,٨٣٣	١,٢٣٩
٨	٢٦,١٢٥	٢٠,٠٩٠	١٥,٥٠٧	١٣,٣٦٢	٧,٣٤٤	٣,٤٩٠	١,٦٤٦
٩	٢٧,٨٧٧	٢١,٦٦٦	١٦,٩١٩	١٤,٦٨٤	٨,٣٤٣	٤,١٦٨	٢,٠٨٨
١٠	٢٩,٥٨٨	٢٣,٢٠٩	١٨,٣٠٧	١٥,٩٨٧	٩,٣٤٢	٤,٨٦٥	٢,٥٥٨
١١	٣١,٢٦٤	٢٤,٧٢٥	١٩,٦٧٥	١٧,٢٧٥	١٠,٣٤١	٥,٥٧٨	٣,٠٥٣
١٢	٣٢,٩٠٩	٢٦,٢١٧	٢١,٠٢٦	١٨,٥٤٩	١١,٣٤٠	٦,٣٠٤	٣,٥٧١
١٣	٣٤,٥٢٨	٢٧,٦٨٨	٢٢,٣٦٢	١٩,٨١٢	١٢,٣٤٠	٧,٠٤٢	٤,١٠٧
١٤	٣٦,١٣٣	٢٩,١٤١	٢٣,٦٨٥	٢١,٠٦٤	١٣,٣٣٩	٧,٧٩٠	٤,٦٦٠
١٥	٣٧,٦٩٧	٣٠,٥٧٨	٢٤,٩٩٦	٢٢,٣٠٧	١٤,٣٣٩	٨,٥٤٧	٥,٢٢٩
١٦	٣٩,٢٥٢	٣٢,٠٠٠	٢٦,٢٩٦	٢٣,٥٤٢	١٥,٣٣٨	٩,٣١٢	٥,٨١٢
١٧	٤٠,٧٩٠	٣٣,٤٠٩	٢٧,٥٨٧	٢٤,٧٦٩	١٦,٣٣٨	١٠,٠٨٥	٦,٤٠٨
١٨	٤٢,٣١٢	٣٤,٨٠٥	٢٨,٨٦٩	٢٥,٩٨٩	١٧,٣٣٨	١٠,٨٦٥	٧,٠١٥
١٩	٤٣,٨٢٠	٣٦,١٩١	٣٠,١٤٤	٢٧,٢٠٤	١٨,٣٣٨	١١,٦٥١	٧,٦٣٣
٢٠	٤٥,٣١٥	٣٧,٥٦٦	٣١,٤١٠	٢٨,٤١٢	١٩,٣٣٧	١٢,٤٤٣	٨,٢٦٠
٢١	٤٦,٧٩٧	٣٨,٩٣٢	٣٢,٦٧١	٢٩,٦١٥	٢٠,٣٣٧	١٣,٢٤٠	٨,٨٩٧
٢٢	٤٨,٢٦٨	٤٠,٢٨٩	٣٣,٩٢٤	٣٠,٨١٣	٢١,٣٣٧	١٤,٠٤١	٩,٥٤٢
٢٣	٤٩,٧٢٨	٤١,٦٣٨	٣٥,١٧٢	٣٢,٠٠٧	٢٢,٣٣٧	١٤,٨٤٨	١٠,١٩٦
٢٤	٥١,١٧٩	٤٢,٩٨٠	٣٦,٤١٥	٣٣,١٩٦	٢٣,٣٣٧	١٥,٦٥٩	١٠,٨٥٦
٢٥	٥٢,٦٢٠	٤٤,٣١٤	٣٧,٦٥٢	٣٤,٣٨٢	٢٤,٣٣٧	١٦,٤٧٣	١١,٥٢٤

فهرست